Titre: Title:	Performance analysis of cooperative MIMO networks
Auteur: Author:	Seyed Mohammad Ali Torabi
Date:	2008
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Torabi, S. M. A. (2008). Performance analysis of cooperative MIMO networks [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/8371/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/8371/
Directeurs de recherche: Advisors:	
Programme: Program:	Non spécifié

UNIVERSITÉ DE MONTRÉAL

PERFORMANCE ANALYSIS OF COOPERATIVE MIMO NETWORKS

SEYED MOHAMMAD ALI TORABI
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE ÉLECTRIQUE)
AOÛT 2008

Your file Votre référence
ISBN: 978-0-494-46084-9
Our file Notre référence
ISBN: 978-0-494-46084-9

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

PERFORMANCE ANALYSIS OF COOPERATIVE MIMO NETWORKS

présenté par: TORABI Seyed Mohammad Ali

en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de:

M. Christian CARDINAL, Ph.D., président

M. Jean-François FRIGON, Ph.D., membre et directeur de recherche

M. David HACCOUN, Ph.D., membre

To: My parents and Laila.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, DR. JEAN-FRANÇOIS FRIGON for his commitment, guidance, encouragement and criticism. He supported me financially and gave me a free hand to explore different research areas from which I learned a lot. I would also like to thank my commitee members, DR. DAVID HACCOUN and DR. CHRISTIAN CARDINAL for their helpful comments and suggestions.

I extend my thanks to my colleagues and friends FARAZ MIRZAEI, ALIREZA RABBANI, DR. MEHDI KHABAZIAN, DR. MOHAMMAD TORABI, YONY-KEVIN AH-KIOON, DR. ZOUHEIR REZKI, DIEGO ENRIQUE PEREA, BEN-WAH KUANG, SHAHRZAD SEYYEDI and MAHMOUD HAGHDOOST for the helpful discussions we had together. I am also very grateful for having HOUMAN ZARRABI as a twenty-year friend and wish to thank him for all the favors he has granted to me and for the great times we had together in Montréal.

Most importantly, I would like to thank my family and especially my parents for their never-ending unconditional love and support, and also for their constant encouragement during my life. Without their devotion and sacrifice, I wouldn't have been able to accomplish my goals. I'm also grateful to my aunt NASTARAN HAGHANI and uncle DR. REZA BATENI for their help, advice and support since my stay in Canada. Last of all but not least of all, I would like to thank my best friend, LAILA SHAMSAIE from whom I got the motivation to continue graduate studies in Canada.

MOHAMMAD ALI TORABI

RÉSUMÉ

Les communications à antennes multiples et les communications coopératives sont parmi les technologies de pointe qui ont été les plus actives dans la recherche académique en communication sans fil au cours des dernières années. L'atout principal de ces deux technologies consiste à exploiter la dimension spatiale facilement disponible. Lorsque ces deux technologies sont combinées ensemble, un gain additionel en puissance peut être obtenu. Un exemple pratique d'une telle combinaison est la communication à antennes multiples par relayage. Dans cette configuration, la communication entre la (les) source(s) et la (les) destination(s) est assisté par un (ou plusieurs) noeud(s) relais qui sont équipés d'antennes multiples.

L'objectif principal de cette thèse est de présenter une technique permettant de choisir les relais et les faisceaux d'antennes appropriés pour des canaux relais à antennes multiples du deux bonds (parallèle ou interférence). Nous démontrons que la technique d'annulation d'interférence pour le canal à relais à deux bonds surpasse l'annulation d'interférence distribuée conventionnelle pour les réseaux à relais pour un nombre pratique de noeuds de relais.

Dans la première partie de la thèse, nous considérons un canal à relais à antennes multiples parallèle à deux bonds où une seule paire de source-destination équipée de M antennes, communique par l'intermédiaire de K relais semi-duplex utilisant la technique d'amplification et relayage (amplify-and-forward) chaqun équipé de N antennes. Une technique de beamforming à multi-usagers est conjointement appliquée à la source et à la destination respectivement afin de multiplexer et de récupérer les données. Afin d'exploiter

d'avantage la diversité multi-relais, nous proposons un nouveau critère de sélection de relais basé sur la *quasi-orthogonalité* des paires de *vecteurs propres (eigenmodes) spatiales* ainsi que sur les paires d'*antennes* des relais. Comparés à des stratégies précédemment proposé pour les canaux à relais à antennes multiples, les algorithmes proposés ont une complexité très faible, diminuent la quantité de rétroaction, et consomment moins de puissance à travers les relais . En outre, pour un nombre de noeuds de relais petit à moyen, la capacité des algorithmes proposés surpassent les autres stratégies et, asymptotiquement lorsque le nombre des noeuds de relais est grand, la capacité des algorithmes proposés augmente comme $\frac{M}{2} \log \log K$ $(K \to \infty)$. Dans la deuxième partie, nous considérons des réseaux d'interférence à relais à antennes multiples à deux-bonds. Plusieurs noeuds de source communiquent avec plusieus noeuds de destination par l'intermédiaire de K noeuds de relais chacun équipé de N antennes. Basé sur des techniques récentes en caractérisation des limites de réseaux X à antenne multiples et le concept de *l'alignement d'interférence*, nous traitons ce réseau en tant que deux réseaux en cascade de X à antennes multiples. De ce fait, nous décomposons le réseau en L canaux non-intéferants à relais à antennes multiples parallèles. Le raisonnement derrière cette approche peut être justifié par la théorie de l'information de canaux X à antennes multiples qui stipule que ces canaux fournissent des degrés de liberté plus élevés qu'un canal à interférence à antennes multiples. La signalisation, l'alignement d'interférence ainsi que le choix de relais sont discutés et des résultats de simulation sont présentés. Comparé à la technique d'annulation d'interférence distribuée, l'approche proposée offre une capacité plus élevée au prix d'une rétroaction de l'état du

canal et au prix d'une signalisation plus sophistiquée.

ABSTRACT

Multiple-antenna and cooperative communications are two advanced technologies that have shaped the realm of academic research on wireless communication over the last few years. The main leverage of these technologies is due to exploiting spatial dimension which is easily available. When multi-antenna and distributed communication are combined together, additional leverages may be gained. Practical examples of such combination is MIMO relay channels when the communication between multi-antenna information source and destination is assisted by a (multiple) relay node(s) which is (are) equipped with multiple antennas.

The main focus of this thesis is on relay selection and beamforming for two-hop MIMO (parallel or interference) relay channels. We show that two-hop interference cancellation technique outperforms conventional distributed interference cancellation for MIMO multi-hop networks for practical number of relay nodes.

In the first part of the thesis, we consider a two-hop MIMO parallel relay channel where a single source-destination pair both equipped with M antennas, communicate via K N-antenna half-duplex amplify-and-forward relays ($K, N \geq M$). A multi-user beamforming technique is jointly applied at the source and destination to respectively multiplex and recover the data streams. In order to further exploit the multi-relay diversity, we propose new relay selection criteria based on *semi-orthogonality* among *spatial eigenmode* and *antenna* pairs of the relays. The proposed algorithms have very low complexity, decrease the amount of feedback, and consume less power across the

relays compared to previously proposed MIMO relaying strategies. Furthermore, for small to medium number of relay nodes, the proposed algorithms' capacity outperform the other strategies and in the asymptote of large number of relay nodes ($K \rightarrow \infty$), the capacity of the proposed algorithms scales as $\frac{M}{2} \log \log K$.

In the second part, we consider two-hop MIMO interference relay networks, where L M-antenna source nodes communicate to L M-antenna destination nodes via K relay nodes each equipped with N antennas. Based on recent advances in characterizing the limits of MIMO X channel and the concept of *interference alignment*, we treat this network as two cascaded MIMO X networks. By doing so, we decompose the network into L *non-interfering* MIMO parallel relay channels. The rationale behind this approach can be justified by information theory of MIMO X channels which reveals that they provide higher degrees of freedom than conventional MIMO interference channels. Signaling, interference alignment as well as relay selection will be discussed and simulation results will be presented. Compared to the well-known *distributed interference cancellation* technique, this approach offers higher capacity at the cost of feeding back channel state information from the relays to the source nodes and more advanced signaling.

CONDENSÉ

L'utilisation des antennes multiples aux deux extrémités d'un lien sans-fil point à point, connue sous le nom de MIMO est une technologie puissante permettant d'améliorer la performance du système. Depuis la fin des années 1990, l'intérêt des systèmes MIMO a beaucoup augmenté de sorte qu'ils sont maintenant déployés dans les réseaux cellulaires sans-fil de la troisième génération (3G) et sont considérés comme la technologie principale des normes de la quatrième génération (4G) des systèmes de communication sans-fil.

Un système de communication MIMO est défini comme un système composé d'un émetteur et d'un récepteur chacun équipé d'antennes multiples. L'avantage des systèmes MIMO provient de l'introduction d'une dimension additionnelle: *la dimension spatiale.* Cette dimension additionnelle permet d'aboutir à une augmentation de l'ordre du gain diversité et des degrés de liberté pour des canaux bien-conditionnés. Un canal bien-conditionné pour un système MIMO fournit des chemins stochastiquement indépendants (non-corrélés) entre l'émetteur et le récepteur. Ces canaux sont observés lorsque:

1. Le niveau de dispersion et de réflexion dans l'environnement (qui sont des phénomènes inhérents des canaux multi-trajets à evanouissement) est élevé.

2. La longueur du réseau des antennes de transmission et de réception est assez longue et l'espacement entre les antennes est plus grand que la

moitié d'une longueur d'onde.

En général, il y a trois gains qui peuvent être réalisés dans des canaux MIMO bien-conditionnés. Ce sont les gains de diversité, de multiplexage et de puissance qui sont utilisés pour réduire la probabilité d'erreur, pour augmenter le taux de transmission et pour augmenter le SNR, respectivement.

Avec des architectures cellulaires conventionnelles, les débits envisagés pour les systèmes sans-fil de la quatrième génération ne semblent pas être atteignables. Même les techniques avancées des systèmes MIMO ne peuvent pas fournir seules assez de puissance pour réaliser le niveau désiré de performance. Dans ce contexte, l'utilisation des techniques de communication distribuée et le *relayage* en particulier, pourraient fournir une puissance additionnelle sans exiger des coûts significatifs à l'infrastructure. Le relayage sans-fil est un outil puissant avec un vaste domaine d'application pour fournir une transmission fiable et un haut débit dans les réseaux sans-fil.

Dans un environnement cellulaire, un relais peut être déployé dans les secteurs où il y a des effets d'ombrage, comme dans les bâtiments et les tunnels. Pour les réseaux ad hoc, le relayage est essentiel non seulement pour surmonter l'effet d'ombrage causé par des obstacles, mais il est aussi nécessaire pour réduire la puissance de la transmission et l'interférence aux noeuds voisins. Le déploiement dynamique des relais est utile pour augmenter la fiabilité du réseau et réduire la probabilité d'erreur dans des applications tactiques. Avec cette motivation, il y a eu récemment un grand intérêt dans les centres de recherches universitaires ainsi que dans l'industrie sur le concept du relayage

pour des réseaux sans-fil (réseaux sans-fil locaux, réseaux cellulaires 4G et réseaux sans-fils à bande large).

Dans ce mémoire, nous considérons un réseau à relais à antennes multiples à deux bonds où la communication entre la (les) source(s) et la (les) destination(s) est assisté par plusieurs noeuds relais qui sont équipés d'antennes multiples. Nous supposons également que les sources et les destinations fonctionnent en mode de multiplexage spatial, c.à.d. des signaux statistiquement indépendants sont transmis par des différentes antennes du (des) noeud(s) de source vers le (les) noeud(s) de destination.

Afin de réduire la complexité du traitment de signal à l'émetteur et au récepteur, nous supposons que tous les noeuds traitent leurs signaux linéairement. En nous basant sur ces hypothèses, nous établissons des algorithmes permettant de choisir les relais et de calculer les faisceaux de transmission et de réception. Ces algorithmes sont de faible-complexité et exigent peu de rétroaction sur l'état du canal. Les résultats de simulation démontrent que les algorithmes proposés surpassent les techniques d'annulation d'interférence distribuées pour de petits ou moyens nombres de noeuds de relais .

Puisque ce mémoire est basé sur le canal MIMO à relais, nous présentons une analyse générale des communications MIMO et des canaux de transmission à relais aux chapitres 2 et 3. Au chapitre 2, nous présentons une analyse théorique de la capacité des systèmes MIMO en posant différentes hypothèses concernant la connaissance du canal. De plus, nous déterminons également

les valeurs limites sur le gain de puissance et sur le gain de diversité des canaux MIMO.

Au chapitre 3, nous présentons le canal à relais composé d'une source d'information, d'une destination d'information et d'un noeud de relais. Divers protocoles de relayage comme Time-Divison Duplex (TDD) et Full-Duplex (FD) sont étudiés. Ainsi plusieurs stratégies de relayage sont analysées comme: amplify-and-forward (AF), decode-and-forward (DF) et compress-and-forward (CF). Enfin nous comparons ces protocoles et ces stratégies de relayage en termes de capacité à différents régimes de SNR.

Au chapitre 4, nous étudions les canaux de relais de parallèle MIMO où il y a une seule paire de source destination et un grand nombre de noeuds de relais. Nous considérons un canal à relais à antennes multiples parallèle à deux bonds où une seule paire de sourcedestination équipée de M antennes, communique par l'intermédiaire de K relais semi-duplex utilisant la technique d'amplification et relayage (amplify-and- forward) chacun équipé de N antennes.

Une technique de beamforming à multi-usagers est conjointement appliquée à la source et à la destination respectivement afin de multiplexer et de récupérer les données. Afin d'exploiter davantage la diversité multi-relais, nous proposons un nouveau critère de sélection de relais basé sur la quasi-orthogonalité des paires de vecteurs propres (eigenmodes) spatiales ainsi que sur les paires d'antennes des relais. Comparés à des stratégies précédemment

proposé pour les canaux à relais à antennes multiples, les algorithmes proposés ont une complexité très faible, diminuent la quantité de rétroaction, et consomment moins de puissance à travers les relais. En outre, pour un nombre petit à moyen de noeuds de relais , la capacité des algorithmes proposés surpassent les autres stratégies et, asymptotiquement lorsque le nombre des noeuds de relais est grand ($K \to \infty$), la capacité des algorithmes proposés augmente comme $\frac{M}{2} \log \log K$.

Au chapitre 5, nous considérons des réseaux d'interférence à relais à antennes multiples à deux-bonds. Plusieurs noeuds de source communiquent avec plusieurs noeuds de destination par l'intermédiaire de K noeuds de relais chacun équipé de N antennes. Basé sur des techniques récentes en caractérisation des limites de réseaux X à antenne multiples et le concept de l'alignement d'interférence, nous traitons ce réseau en tant que deux réseaux X en cascade à antennes multiples.

De ce fait, nous décomposons le réseau en L canaux non-interférents à relais à antennes multiples parallèles. Le raisonnement derrière cette approche peut être justifié par la théorie de l'information de canaux X à antennes multiples qui stipule que ces canaux fournissent des degrés de liberté plus élevés qu'un canal à interférence à antennes multiples. La signalisation, l'alignement d'interférence ainsi que le choix de relais sont discutés et des résultats de simulation sont présentés. Comparé à la technique d'annulation d'interférence distribuée, l'approche proposée offre une capacité plus élevée au prix d'une rétroaction de l'état du canal et d'une signalisation plus sophistiquée.

Finalement, au chapitre 6, nous présentons des conclusions et nous suggérons des portes de recherche basées sur les idées développées dans ce mémoire.

TABLE OF CONTENTS

LIST OF FIGURES

NOTATIONS

3G	Third Generation
4G	Forth Generation
AF	Amplify-and-Forward
BC	Broadcast Channel
BF	Beamforming
CF	Compress-and-Forward
CI	Channel Inversion
CDF	Cumulative Density Function
CSI	Channel State Information
CSIR	Channel State Information at Receiver
CSIT	Channel State Information at Transmitter
DF	Decode-and-Forward
DOF	Degrees of Freedom
DPC	Dirty Paper Coding
FD	Full Duplex
FDMA	Frequency-Division Multiple Access
HD	Half Duplex
IC	Interference Channel
i.i.d.	Independently Identically Distributed
ISI	Inter-Stream Interference
IUI	Inter-User Interference
LoS	Line of Sight
MIRN	MIMO Interference Relay Network

MISO	Multiple-Input Single-Output
MIMO	Multiple-Input Multiple-Output
MU	Multi User
MMSE	Minimum Mean Square Error
p.d.f.	Probability Density Function
QoS	Quality of Service
SNR	Signal to Noise Ratio
SINR	Signal to Interference plus Noise Ratio
TDD	Time Division Duplex
TDMA	Time-Division Multiple Access
ZF	Zero Forcing
ZMCSCG	Zero-Mean Circularly-Symmetric Complex Gaussian
$\mathbb{C}$	Complex Field
$\mathbb{R}$	Real Field
$\mathbb{Z}^+$	Positive Integer Field
$\mathbf{A}$	Bold Capital Letters Denote Matrices
$\mathbf{a}$	Bold Small Letters Denote Vectors
$\mathbf{I}_n$	Identity Matrix of Dimension n
$Tr(\mathbf{A})$	Trace of Matrix $\mathbf{A}$
$\det(\mathbf{A})$	Determinant of Squared Matrix $\mathbf{A}$
$\mathbf{A}^T$	Transpose of Matrix $\mathbf{A}$
$\mathbf{A}^*$	Conjugate Transpose (Hermitian) of Matrix $\mathbf{A}$ with Complex Entries
$\mathbf{A}^\dagger$	Pseudo-inverse of Matrix $\mathbf{A}$
$\mathbb{N}(\boldsymbol{A})$	Null-Space of Matrix $\boldsymbol{A}$

$\mathbb{E}\{X\}$	Expectation over Random Variable X
$\mathbb{P}\{\mathcal{A}\}$	Probability of Event $\mathcal{A}$
$\mathbb{OC}^{M\times N}$	The set of All $M \times N$ Complex Matrices with Mutually Orthogonal and Normal Columns
$\Omega(\boldsymbol{A})$	Sub-Space Spanned By Columns of $\boldsymbol{A}$
$\doteq$	Asymptotically Equivalent
$u(x) = \mathcal{O}(v(x))$	$\lim_{x\to\infty} \frac{u(x)}{v(x)} = 1$
$\boldsymbol{A} \perp \boldsymbol{B}$	Each column of matrix $\boldsymbol{A}$ is Orthogonal to all columns of $\boldsymbol{B}$

LIST OF TABLES

CHAPTER 1

INTRODUCTION AND BACKGROUND WORK

1.1 Introduction

The use of multiple antennas at both ends of a point-to-point wireless link, known as multiple-input multiple-output (MIMO) wireless, is a powerful performance enhancing technology [1–4]. In point-to-point wireless links, MIMO systems improve spectral efficiency, link reliability and power efficiency through spatial multiplexing gain, diversity gain and array gain[1], respectively.

The very high data rates envisioned for fourth-generation (4G) wireless systems in reasonably large or dense areas do not appear to be feasible with conventional cellular architectures. Multiple antenna techniques and advanced signal processing techniques (such as interference cancellation algorithms) by themselves cannot provide enough leverage to achieve the desired level of performance. In this context, *distributed* communication techniques and *relaying* in particular, could provide an additional leverage without requiring significant infrastructure deployment costs. Wireless relaying is a powerful tool to provide reliable transmission, high throughput and broad coverage for wireless networks in a variety of applications. In a cellular environment,

[1]These three terms are defined in section 2.1.

a relay can be deployed in areas where there are strong shadowing effects, such as inside buildings and tunnels. For mobile ad hoc networks, relaying is essential not only to overcome shadowing due to obstacles but also to reduce unnecessary transmission power and hence radio frequency interference to neighboring nodes. For tactical applications, dynamic deployment of manned or unmanned relays are useful to enhance the network's reliability, throughput and low probability of detection and/or interception.

With this motivation, there has recently been growing interest both in academia and industry in the concept of relaying in infrastructure-based wireless networks such as next generation cellular (B3G, 4G), wireless local area networks (WLANs) (802.11) and broadband fixed wireless (802.16) networks [5]. Distributed cooperation techniques over single-antenna relay channels have been studied by [6–8] and extended to the MIMO relay channel in [9]. Capacity scaling laws and Shannon-theoretic power-bandwidth trade-offs over large-scale MIMO relay networks were characterized in [10–13], respectively.

In this thesis, we consider a two-hop fading MIMO interference relay network (MIRN), where the communication between the multi-antenna source node(s) and destination node(s) is assisted by several multi-antenna relays between the source(s) and destination(s). We also assume that the source(s) and destination(s) operate in spatial multiplexing mode, i.e. statistically independent signals are transmitted from the different antennas of the source node(s) toward the destination node(s). In order to reduce the transceiver complexity, we constrain all nodes to process their signals linearly. Based on these assumptions, we design low-complexity relay selection and beam-

forming algorithms that require small amount of channel state information feedback. Simulation results show that the proposed algorithms outperform distributed interference cancellation techniques for small to medium number of relay nodes.

1.2 Contributions

This thesis is concerned with information theoretic analyses of MIMO parallel relay channels and MIMO interference relay networks. The main contributions contained in this study can be classified into two broad categories:

1. We consider a two-hop MIMO parallel relay channel where a single source-destination pair both equipped with M antennas, communicate via K N-antenna half-duplex amplify-and-forward relays ($K, N \geq M$). Multi-user beamforming techniques are jointly applied at the source and destination to respectively multiplex and recover the data streams. In order to effectively exploit the multi-relay diversity, we propose new relay selection criteria based on *semi-orthogonality* among *spatial eigenmode* and *antenna* pairs of the relays. The proposed algorithms have very low complexity, decrease the amount of feedback, and consume less power across the relays compared to previously proposed MIMO relaying strategies. Furthermore, for small to medium number of relay nodes, the proposed algorithms' capacity outperform the other strategies and in the asymptote of large number of relay nodes ($K \to \infty$), the capacity of the proposed algorithms scales as $\frac{M}{2} \log \log K$ [14].

2. We consider a class of wireless ad hoc networks, namely *two-hop MIMO interference relay networks*, where L M-antenna source nodes communicate to L M-antenna destination nodes via K relay nodes each equipped with N antennas. Based on recent advances in characterizing the limits of MIMO X channel and the concept of *interference alignment*, we treat this network as two cascaded MIMO X networks. By doing so, we decompose the network into L *non-interfering* parallel MIMO relay channels. Signaling, interference alignment, relay selection will be discussed and simulation results will be presented. Compared to the well-known *distributed interference cancellation* technique, this approach offers higher capacity at the cost of feeding back CSI from the relays to the source nodes and more advanced signaling [15].

1.3 Thesis Outline

Since MIMO relay channels form the backbone of this thesis, we present detailed analysis of MIMO communications and relay channels in chapters 2 and 3 respectively. In chapter 2, we investigate the information theoretic capacity analysis of MIMO systems with different channel knowledge assumptions. Moreover, bounds on the array gain and diversity gain of MIMO channels in the so-called "beamforming (BF)" mode are derived.

In chapter 3, we introduce the relay channel consisting of an information source, an information destination and a relay node. Various relaying protocols, namely full duplex (FD), time-division duplex (TDD) and half duplex

(HD) are investigated as well as several relaying strategies such as: amplify-and-forward (AF), decode-and-forward (DF), and compress-and-forward (CF). We compare these relaying protocols and strategies in terms of capacity at different SNR regimes.

In chapter 4, we study the MIMO parallel relay channels where there is a single source-destination pair and a large number of relay nodes. Semi-orthogonal relay selection and beamforming at the source and destination is discussed and the system capacity scaling is presented.

In chapter 5, we study the MIMO interference relay network (MIRN) where multiple source-destination communicate via multiple relay nodes which are located between the source and destination nodes. Based on the concept of "interference alignment", we consider the first and second hop as two cascaded MIMO X channel and decompose the MIRN into multiple non-interfering MIMO parallel relay channels as described in chapter 4. Signaling, relay selection and interference alignment is discussed throughout this chapter.

Finally, in chapter 6, we provide concluding remarks and discuss possible future works based on the ideas developed in this thesis.

CHAPTER 2

INTRODUCTION TO MIMO COMMUNICATIONS

2.1 Preliminaries

Wireless communication strives for higher *data rates* as well as better *quality of service* (QoS). The use of multiple-input multiple-output (MIMO) communication systems is aimed at these goals [1]. Since the late 90's, interest in MIMO systems has exploded such that they are now deployed in third-generation (3G) wireless cellular networks and are considered as a key technology for next-generation wireless communication standards (4G). A MIMO communication system is defined as a system of transmitter and receiver each equipped with multiple antennas. The significance of MIMO systems stems from the introduction of an additional dimension: the *spatial dimension.* This additional dimension can lead to an increase in the *diversity order* and in the *degrees of freedom* [16] over well-conditioned channels. Well-conditioned channels for MIMO systems are the ones that provide stochastically independent (uncorrelated) signal paths from the transmitter to the receiver. Such channels are observed when [17]:

1. The amount of scattering and reflection in the environment (which is an inherent part of multi-path fading channels) is large.

2. The length of the transmit and receive antenna arrays is long enough

and the antenna spacing is larger than half a wavelength.

In this case, the surprising feature of MIMO systems appears. *Multi-path fading* [18], which was known to be detrimental for wireless communications, has ironically turned out to be an advantage for MIMO systems. In other words, multi-path fading increases the degrees of freedom of the system which may be used to:

1. increase data rate via *spatial multiplexing* [4],
2. reduce bit error rate via *space-time coding* [3, 19] or
3. a combination of both [20–22].

In general, there are three gains that can be achieved in MIMO systems and any MIMO configuration may be described by the amount of those gains. They are the *diversity, multiplexing* and *array gains.* Due to frequent use of these terms throughout this thesis, their definitions are mentioned according to [23]:

1. *Diversity Gain*: The gain provided by the spatially independent links either at the transmitter, the receiver or both, and is necessary for mitigating the adverse effects of fading.
2. *Multiplexing Gain*: Capacity gain at no additional power or bandwidth consumption obtained through the use of multiple antennas at both sides of a wireless link which increases the number of degrees of freedom available to send data.

3. *Array Gain*: The average increase in the SNR at the receiver that arises from the coherent combining effects of multiple antennas at the transmitter, receiver or both. This is the gain provided by the channel state information (CSI).

In the rest of this chapter, we present the basic information-theoretic data rates supported by MIMO systems under different CSI assumptions. We will also discuss the bounds on the diversity and array gains supported by MIMO systems. In section 2.2, a probabilistic model for MIMO channel is introduced. Section 2.3 treats two important notions for the capacity characterization in wireless channels: ergodic capacity and outage capacity. In this section, we derive the ergodic capacity for two channel knowledge assumptions (CSIR and CSIT). Finally in section 2.4, we will consider a special MIMO configuration called "beamforming" or "maximum ratio transmission" and will derive the bounds on array and diversity gains achieved by this configuration. In each section, numerical results are presented to give further insights into theoretical aspects.

2.2 MIMO System Model

We consider a MIMO system with M_t transmit antennas and M_r receive antennas as shown in Figure 2.1. The MIMO transmitter produces a channel input symbol $\mathbf{x}$, which is an $(M_t \times 1)$ vector over the complex field, per channel use. The corresponding channel output symbol is given by $\mathbf{y}$ which is a $(M_r \times 1)$ dimensional vector over the complex field. The probabilis-

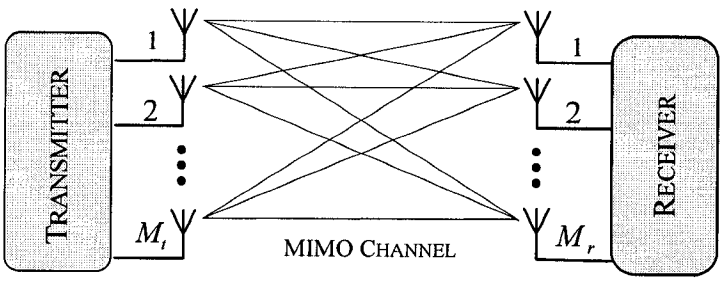

Figure 2.1 Block diagram of an $(M_r \times M_t)$ MIMO system.

tic MIMO channel can be characterized by $(M_r \times M_t)$-dimensional channel state matrix $\mathbf{H}$ which is over the complex field. In general, each channel state realization $\mathbf{H} = \boldsymbol{\mathcal{H}}$, specifies a channel state transition probability $\mathbb{P}(\mathbf{y} = \boldsymbol{y}|\mathbf{x} = \boldsymbol{x}, \mathbf{H} = \boldsymbol{\mathcal{H}})$. Without loss of generality, we assume discrete inputs and discrete outputs. Generalization to continuous inputs and continuous outputs is straight-forward.

The channel transition probability of MIMO channels can be completely characterized by the conditional channel transition probability and the channel state sequence probability as

$$\mathbb{P}(\boldsymbol{y}_1^N|\boldsymbol{x}_1^N) = \sum_{\boldsymbol{\mathcal{H}}_1^n} \mathbb{P}(\boldsymbol{y}_1^N|\boldsymbol{x}_1^N, \boldsymbol{\mathcal{H}}_1^N)\mathbb{P}(\boldsymbol{\mathcal{H}}_1^N), \tag{2.1}$$

where $\boldsymbol{x}_1^N = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]$ denotes a block of N transmitted symbols, $\boldsymbol{y}_1^N = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N]$ denotes a block of N received symbols and $\boldsymbol{\mathcal{H}}_1^N = [\boldsymbol{\mathcal{H}}_1, \ldots, \boldsymbol{\mathcal{H}}_N]$

denotes a block of N channel states. In general, the output symbol at any time depends not only on the current transmitted symbol but also on the past and future transmitted symbols. If this is the case, the channel is said to have *memory*. Otherwise, the probabilistic channel is said to be *memoryless*. In other word a channel is said to be memoryless if the unconditional transitional probability can be expressed into a product form:

$$\mathbb{P}(\boldsymbol{y}_1^N|\boldsymbol{x}_1^N) = \prod_{n=1}^{N} \mathbb{P}(\boldsymbol{y}_n|\boldsymbol{x}_n) \tag{2.2}$$

Now we look at two important channel models.

2.2.1 Fast Flat Fading MIMO Channels

The discrete-time input-output relationship is given by:

$$\boldsymbol{y}_n = \boldsymbol{\mathcal{H}}_n\boldsymbol{x}_n + \boldsymbol{z}_n \tag{2.3}$$

where $\boldsymbol{z}_n$ is a $(M_r \times 1)$-dimensional independently identically distributed (i.i.d) zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise vector with unit variance. In this model, $\boldsymbol{x}_n$ and $\boldsymbol{y}_n$ are the input and output complex vectors. $\{\boldsymbol{\mathcal{H}}_n\}$ are i.i.d. complex channel state sequences. Given $\boldsymbol{\mathcal{H}}_1^N$ and $\boldsymbol{x}_1^N$, the channel transition probability is given by

$$\mathbb{P}(\boldsymbol{y}_1^N|\boldsymbol{x}_1^N, \boldsymbol{\mathcal{H}}_1^N) = \prod_{n=1}^{N} \mathbb{P}(\boldsymbol{y}_n|\boldsymbol{x}_n, \boldsymbol{\mathcal{H}}_n) \tag{2.4}$$

Together with the i.i.d. channel state sequence property:

$$\mathbb{P}(\boldsymbol{\mathcal{H}}_1^N) = \prod_{n=1}^{N} \mathbb{P}(\boldsymbol{\mathcal{H}}_n), \tag{2.5}$$

the fast flat fading channel is equivalent to a *memoryless* channel because the channel transition probability can be decomposed into a product form.

2.2.2 Quasi-static Flat Fading MIMO Channels

As the transmission bit rate increases and the frame duration becomes shorter, we may have a *slow-fading* situation across the entire symbols of a frame. In fact, the transmit symbols over a frame will share the same channel fading; that is, $\boldsymbol{\mathcal{H}}_1 = \boldsymbol{\mathcal{H}}_2, \ldots = \boldsymbol{\mathcal{H}}_N = \boldsymbol{\mathcal{H}}$. Such slow-fading situation is called *quasi-static fading*. The quasi-static flat-fading MIMO channel has the transition probability $\mathbb{P}(\boldsymbol{y}_1^N|\boldsymbol{x}_1^N, \boldsymbol{\mathcal{H}}_1^N) = \prod_n \mathbb{P}(\boldsymbol{y}_n|\boldsymbol{x}_n, \boldsymbol{\mathcal{H}})$. Because of quasi-fading channel across entire frame, the quasi-static fading channel is *not* memoryless because the unconditional channel transition probability can not be decomposed into the required product form. Hence, this channel is a *non-ergodic memory* channel.

2.3 Capacity Definitions

Now, we shall clarify two important concepts of Shannon's capacity for probabilistic channels with states. For ergodic channels, the interpretation of the

channel capacity is straightforward. Specifically, for any rate $R < C$ there exists at least one encoder and decoder that achieves arbitrarily small probability of error. For non-ergodic channels, this is no longer the case because channel capacity in Shannon's sense is zero. In general, we have two important definitions of channel capacity, namely the *ergodic capacity* and the *outage capacity* which are described in the sequel.

2.3.1 Ergodic Capacity

Ergodic capacity refers to the channel capacity in Shannon's sense; that is, for any rate $R < C$, there exists at least one encoder and one decoder that achieves a small probability of error. Conversely, if $R > C$, the error probability is greater than zero for any encoder and decoder. For example consider the transmission of a block of N symbols $\boldsymbol{x}_1^N = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$ with N symbols across a flat fading channel with state sequence $\boldsymbol{\mathcal{H}}_1^N = \{\boldsymbol{\mathcal{H}}_1, \boldsymbol{\mathcal{H}}_2, \ldots, \boldsymbol{\mathcal{H}}_N\}$. If the encoding frame is large enough (larger than the coherence time of the fading channel), such that $\{\boldsymbol{\mathcal{H}}_1, \boldsymbol{\mathcal{H}}_2, \ldots, \boldsymbol{\mathcal{H}}_N\}$ spans across an *ergodic realization* of ergodic fading process $\mathbf{H}(t)$, the resulting ergodic capacity is nonzero and is given by :

$$C = \lim_{N\to\infty} \frac{1}{N} I(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N; \boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N, \boldsymbol{\mathcal{H}}_1, \boldsymbol{\mathcal{H}}_2, \ldots, \boldsymbol{\mathcal{H}}_N) \qquad (2.6)$$

The ergodic capacity reduces to the following well-known formula:

$$C = I(\mathbf{x}; \mathbf{y}, \mathbf{H}) = I(\mathbf{x}; \mathbf{y}|\mathbf{H}). \qquad (2.7)$$

Now, if the transmitted rate is lower than the ergodic capacity, the error probability is exponentially decaying with the frame length N for capacity achieving codes.

2.3.2 Outage Capacity

The ergodic condition is not necessarily satisfied in many practical communication systems operating over fading channels. In this case, there is no significant channel variation across the encoding frame and there is no classical Shannon meaning to the capacity in this typical situation. In the extreme case, when the frame is very short or the coherence time is very long, the entire encoding frame share a single fading state realization, which is $\boldsymbol{\mathcal{H}}_1 = \boldsymbol{\mathcal{H}}_2 = \ldots = \boldsymbol{\mathcal{H}}_N = \boldsymbol{\mathcal{H}}$. In this case, the channel belongs to the type of *non-ergodic* memory channel and the mutual information becomes *information unstable.* The channel capacity in Shannon's sense is given by

$$C = \liminf \frac{1}{N} I(\boldsymbol{x}_1^N; \boldsymbol{y}_1^N, \boldsymbol{\mathcal{H}}) \tag{2.8}$$

where lim inf is in probability sense and $\frac{1}{N} I(\boldsymbol{x}_1^N; \boldsymbol{y}_1^N, \boldsymbol{\mathcal{H}})$ is the normalized sequence of mutual information. In fact, the mutual information is considered as a function of the random realization of $\boldsymbol{\mathcal{H}}$ and therefore is a random variable itself. Hence, there may be a non-negligible probability that the value of transmission rate exceeds the current capacity. This situation gives rise to error probabilities that do not decay with the increase of block length N. Hence, the ergodic capacity of the channel is zero, meaning that no matter

how small the transmission rate is, there is no guarantee that the transmitted frame will be error-free. Instead of looking at the ergodic capacity in Shannon's sense, we can characterize the capacity from an *outage* perspective. The outage capacity C_{out} at a given outage probability $\mathbb{P}_{out}$ is defined as the maximum data rate R such that

$$\mathbb{P}\{R < C\} \leq \mathbb{P}_{out}. \tag{2.9}$$

In other words, the outage probability $\mathbb{P}_{out}$ is the cumulative distribution function of the random mutual information.

2.3.3 Perfect CSIR

Now we consider the case when we have perfect channel state information at the receiver but have no channel state information at the transmitter side (perfect CSIR but no CSIT). In this case, the MIMO channel has the following transition probability:

$$\mathbb{P}(\boldsymbol{y}_1^N, \boldsymbol{\mathcal{H}}_1^N | \boldsymbol{x}_1^N) = \mathbb{P}(\boldsymbol{y}_1^N | \boldsymbol{\mathcal{H}}_1^N, \boldsymbol{x}_1^N) \mathbb{P}(\boldsymbol{\mathcal{H}}_1^N). \tag{2.10}$$

For ergodic CSI sequences, the channel capacity is given by:

$$C_{CSIR} = \lim_{N\to\infty} \frac{1}{N} \max_{\mathbb{P}(\boldsymbol{x}_1^N)} I(\boldsymbol{x}_1^N; \boldsymbol{y}_1^N, \boldsymbol{\mathcal{H}}_1^N) \tag{2.11}$$

$$= \lim_{N\to\infty} \frac{1}{N} \max_{\mathbb{P}(\boldsymbol{x}_1^N)} I(\boldsymbol{x}_1^N; \boldsymbol{y}_1^N | \boldsymbol{\mathcal{H}}_1^N) + I(\boldsymbol{x}_1^N; \boldsymbol{\mathcal{H}}_1^N) \tag{2.12}$$

$$= \lim_{N\to\infty} \frac{1}{N} \max_{\mathbb{P}(\boldsymbol{x}_1^N)} I(\boldsymbol{x}_1^N; \boldsymbol{y}_1^N | \boldsymbol{\mathcal{H}}_1^N) \tag{2.13}$$

where $I(\boldsymbol{x}_1^N; \boldsymbol{\mathcal{H}}_1^N) = 0$ because the transmitted symbol is independent of the channel state. Furthermore, if the conditional channel transition probability can be decomposed into product form and channel state sequence $\boldsymbol{\mathcal{H}}_1^N$ is i.i.d. then the equivalent channel is memoryless. Hence the channel capacity becomes:

$$C_{CSIR} = \max I(\mathbf{x}; \mathbf{y}|\mathbf{H}). \tag{2.14}$$

In order to derive the ergodic capacity of fast flat fading channels, we write the input-output relationship:

$$\boldsymbol{y}_n = \boldsymbol{\mathcal{H}}_n \boldsymbol{x}_n + \boldsymbol{z}_n. \tag{2.15}$$

From probability theory we know that if $\boldsymbol{x}_n$ and $\boldsymbol{z}_n$ are two independent circular symmetric complex gaussian random vectors, then so are $\boldsymbol{\mathcal{H}}_n \boldsymbol{x}_n$ and $\boldsymbol{y}_n$. Besides that, for a zero-mean random vector $\mathbf{x}$ with covariance matrix $\boldsymbol{R}_{\mathbf{x}}$, the entropy is upper bounded by:

$$h(\mathbf{x}) \leq \log\det(\pi e \boldsymbol{R}_{\mathbf{x}}) \tag{2.16}$$

where equality holds if $\mathbf{x}$ is circularly symmetric complex Gaussian ($h(.)$ denotes entropy) [24]. From information theory, we also know that mutual information becomes

$$I(\mathbf{x};\mathbf{y}|\mathbf{H}) = h(\mathbf{y}|\mathbf{H}) - h(\mathbf{y}|\mathbf{x},\mathbf{H}). \tag{2.17}$$

From (2.16) and (2.17), we then have that the first entropy is upper bounded by $\log\det(\pi e\boldsymbol{R}_{\mathbf{y}})$ while the second entropy is upper bounded by $\log\det(\pi e\boldsymbol{I})$. Furthermore there is a total power constraint on transmit power:

$$Tr(\boldsymbol{R}_{\mathbf{x}}) \leq P. \tag{2.18}$$

Therefore, the capacity of the fast flat fading MIMO channel can be written as:

$$\begin{aligned} C_{CSIR} &= \max\ \mathbb{E}_{\mathbf{H}}\Big[\log\det\Big(\boldsymbol{I}_{M_r} + \mathbf{H}\boldsymbol{R}_{\mathbf{x}}\mathbf{H}^*\Big)\Big] \\ &\quad s.t.\ Tr(\boldsymbol{R}_{\mathbf{x}}) \leq P \end{aligned} \tag{2.19}$$

where the optimization is performed over all input covariance matrices. It can be easily proven that the optimal transmit covariance matrix that achieves the capacity when there is no CSIT is the diagonal matrix that allocates power equally across the transmit antennas [4]. Therefore, we can write the capacity as:

$$C_{CSIR} = \mathbb{E}_{\mathbf{H}}\Big[\log\det\Big(\boldsymbol{I}_{M_r} + \frac{P}{M_t}\mathbf{H}\mathbf{H}^*\Big)\Big]. \tag{2.20}$$

If $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$ are the ordered singular values of $\mathcal{H}$, then we can write 2.20 as

$$\begin{aligned} C_{CSIR} &= \mathbb{E}\Big[\sum_{i=1}^{M} \log\Big(1+\frac{P}{M_t}\lambda_i^2\Big)\Big] \\ &= \sum_{i=1}^{M} \mathbb{E}\Big[\log\Big(1+\frac{P}{M_t}\lambda_i^2\Big)\Big], \end{aligned} \tag{2.21}$$

where $M = \min(M_t, M_r)$. By applying Jensen's inequality [24] to the above formula, we obtain the following inequality:

$$\sum_{i=1}^{M} \log\Big(1+\frac{P}{M_t}\lambda_i^2\Big) \leq M \log\Big(1+\frac{P}{M_t}\Big[\frac{1}{M}\sum_{i=1}^{M}\lambda_i^2\Big]\Big) \tag{2.22}$$

with equality if and only if the singular values are all equal. At high SNR regime, the capacity for the i.i.d. Rayleigh channel is given by

$$C_{CSIR} \approx M \log\frac{P}{M_t} + \sum_{i=1}^{M_t} \mathbb{E}[\log\lambda_i^2], \tag{2.23}$$

and

$$\mathbb{E}[\log\lambda_i^2] > -\infty, \tag{2.24}$$

for all i. Hence, M degrees of freedom is achieved. In fact, we have

$$\sum_{i=1}^{M} \mathbb{E}[\log\lambda_i^2] = \sum_{i=|M_t-M_r|+1}^{\max(M_r,M_t)} \mathbb{E}[\log\chi_{2i}^2] \tag{2.25}$$

where χ_{2i}^2 is χ-squared random variable with $2i$ degrees of freedom.

At low SNR regime, we use the approximation $\log_2(1+x) \approx x \log_2 e$ for $x << 1$ and obtain

$$\begin{aligned} C_{CSIR} &= \sum_{i=1}^{M} \mathbb{E}\Big[\log\Big(1+\frac{P}{M_t}\lambda_i^2\Big)\Big] \\ &\approx \sum_{i=1}^{M} \frac{P}{M_t}\mathbb{E}[\lambda_i^2]\log_2 e \\ &= \frac{P}{M_t}.\mathbb{E}\Big[\sum_{i,j}|h_{ij}^2|\Big]\log_2 e \\ &= M_r.P.\log_2 e. \end{aligned} \tag{2.26}$$

Thus at low SNR, an $M_r \times M_t$ MIMO system yields an array (power) gain of M_r over a SISO system. Note that increasing the number of transmit antenna does no increase the power gain, since unlike the case when there is CSIT, transmit beamforming (or maximum ratio transmission which is discussed in this chapter) can not be done to add signals constructively at the transmitter.

Figure 2.2 shows the capacity of the i.i.d Rayleigh channel ($\boldsymbol{\mathcal{H}} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I})$)for various number of antennas. As we see, at moderate to high SNR, the capacity of an $M \times M$ is about M times the capacity of a SISO channel. The asymptotic slope of capacity versus SNR in dB scales proportional to $M \log \mathsf{SNR}$. In this regime, full degrees of freedom is attained and the impact of array gain is less impressive than that of multiplexing gain. At low SNR, a system with $M \times M$ antenna yields a power gain of M over SISO channels not because of the number of transmit antennas but in fact because of the number of receive antennas which provide a power boost over SISO channels

with no CSIT. In fact, in this regime, there is only one degree of freedom and the impact of array gain is much more than the multiplexing gain.

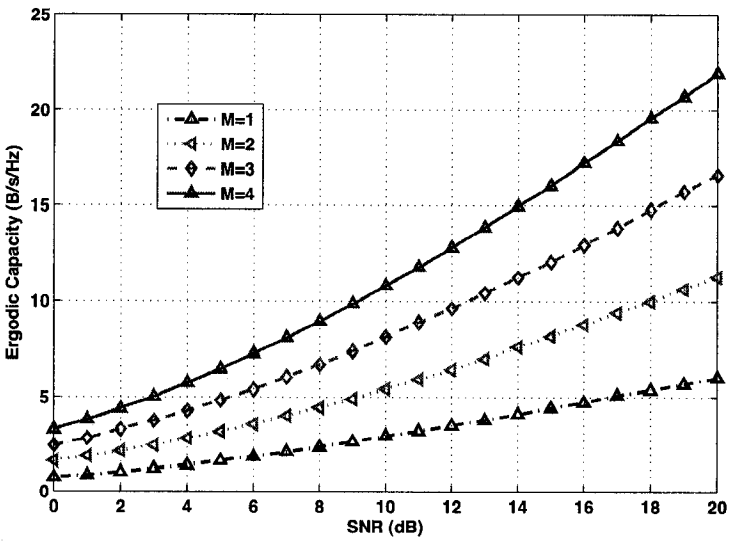

Figure 2.2 Ergodic capacity with perfect CSIR as a function of SNR for $M_t = M_r = M = 1, 2, 3, 4$.

2.3.4 Perfect CSIR and Perfect CSIT

With perfect CSI, *adaptation* in channel encoding is possible and with perfect CSIR channel matched decoding is possible. Hence the capacity of the channels with perfect CSIR and CSIT is higher than the capacity with only perfect CSIR. The MIMO channel capacity with perfect CSIR and CSIT is defined as:

$$C_{CSIR,CSIT} = \mathbb{E}_{\mathbf{H}}\Big[\max_{\mathbb{P}(\mathbf{x}|\mathbf{H}=\mathcal{H})} I(\mathbf{x};\mathbf{y}|\mathbf{H}=\mathcal{H})\Big]. \tag{2.27}$$

For an in-depth analysis of capacity with CSIT , interested readers may refer to [25] for SISO channels and [4, 26] for MIMO channels. As before, we consider a flat fading MIMO channel with input-output relationship $\boldsymbol{y}_n = \boldsymbol{h}_n\boldsymbol{x}_n + \boldsymbol{z}_n$. The ergodic capacity is given by:

$$\begin{aligned} C_{CSIR,CSIT} &= \mathbb{E}_{\mathbf{H}}\Big(\max_{\mathbb{P}(\mathbf{x}|\mathbf{H}=\boldsymbol{\mathcal{H}})} I(\mathbf{x};\mathbf{y}|\mathbf{H}=\boldsymbol{\mathcal{H}})\Big) && (2.28)\\ &= \mathbb{E}_{\mathbf{H}}\Big(\max_{\mathbb{P}(\mathbf{x}|\mathbf{H}=\boldsymbol{\mathcal{H}})} \Big[h(\mathbf{y}|\mathbf{H}=\boldsymbol{\mathcal{H}}) - h(\mathbf{y}|\mathbf{x},\mathbf{H}=\boldsymbol{\mathcal{H}})\Big]\Big) && (2.29)\\ &= \mathbb{E}_{\mathbf{H}}\Big(\max_{\mathbb{P}(\mathbf{x}|\mathbf{H}=\boldsymbol{\mathcal{H}})} h(\mathbf{y}|\mathbf{H}=\boldsymbol{\mathcal{H}})\Big) - h(\mathbf{z}). && (2.30) \end{aligned}$$

Since the channel noise is Gaussian, the capacity achieving distribution is complex gaussian with covariance $\boldsymbol{R}_{\mathbf{x}}(\boldsymbol{\mathcal{H}}) = \mathbb{E}\{\mathbf{x}\mathbf{x}^*|\boldsymbol{\mathcal{H}}\}$. The ergodic capacity of fast flat fading MIMO channels with full CSIT and CSIR is given by:

$$C_{CSIR,CSIT} = \mathbb{E}_{\mathbf{H}}\Big[\max\ \log\det\Big(\boldsymbol{I}_{M_r} + \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{R}}_{\mathbf{x}}\boldsymbol{\mathcal{H}}^*\Big)\Big]. \qquad (2.31)$$

where the input symbols must satisfy an average power constraint

$$\mathbb{E}\Big[Tr(\boldsymbol{R}_{\mathbf{x}})\Big] \leq P. \qquad (2.32)$$

Given any fading channel realization $\boldsymbol{\mathcal{H}}$, the optimum input covariance matrix is [17]:

$$\boldsymbol{R}_{\mathbf{x}}(\boldsymbol{\mathcal{H}}) = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^*, \qquad (2.33)$$

where $\boldsymbol{U}$ is a $M_t \times r$ (r=rank of channel matrix) eigenvector matrix of $\boldsymbol{\mathcal{H}\mathcal{H}}^* = \boldsymbol{U\Lambda U}^*$ and $\boldsymbol{D}$ is a $r \times r$ diagonal matrix with diagonal elements given by:

$$d_{ii} = \left(\frac{1}{\mu} - \frac{1}{\lambda_i^2}\right)^+ \quad i = 1, 2, \ldots, r. \tag{2.34}$$

where the λ_i's are the diagonal elements of $\boldsymbol{\Lambda}$ and μ is a constant that satisfies $\sum_{i=1}^{r} d_{ii} = P$. Here we need no adaptation to achieve ergodic capacity. As it can be inferred, the optimal solution involves both *spatial power water-filling* and *temporal power water-filling*. Spatial water-filling refers to the optimal distribution of power across r spatial channels. More power is allocated to spatial channels having a better condition and less power is allocated to spatial channels that have poor conditions. Temporal power water-filling refers to allocation of total power across various fading symbols temporally. Smaller total power is allocated during time-instances that $\sum_i 1/\lambda_i^2$ is large so as to save power for better time instances. It is easy to prove that for i.i.d. Rayleigh channels[1] at high SNR, the waterfilling solution allocates equal power to all spatial eigenmodes as well as equal amount of power over time. Thus,

$$C_{CSIR,CSIT} \approx \sum_{i=1}^{M} \mathbb{E}\Big[\log(1 + \frac{P}{M}\lambda_i^2\Big]. \tag{2.35}$$

If we compare this with C_{CSIR}, we see that the number of degrees of freedom is unchanged but there is an array gain of M_t/M. Thus when the number of transmit antenna is larger than the number of receive antennas, there is an array gain of M_t/M due to CSIT. At low SNR, there is a further gain from

[1]If $\boldsymbol{\mathcal{H}}$ is an i.i.d. Rayleigh channel matrix, then with probability one, the random channel $\boldsymbol{\mathcal{H}\mathcal{H}}^*$ has full rank.

from CSIT compared to CSIR case due to dynamic allocation of power which gives more power to the stronger eigenmodes. In Figure 2.3 we have plotted the ergodic capacity of a MIMO system with and without CSIT for $M = 2, 4$ antenna as a function of SNR. We see that at low SNR, array gain is the dominant factor and due to dynamic allocation of power, there is an extra power gain compared to the no CSIT case. At high SNR, ergodic capacity with and without CSIT converge, since there is equal number of transmit and receive antennas which results in no gain for $C_{CSIT,CSIR}$ compared to C_{CSIR}. However, if there is more antennas at the transmitter than at the receiver, $C_{CSIT,CSIR}$ will outperform C_{CSIR} at high SNR regime.

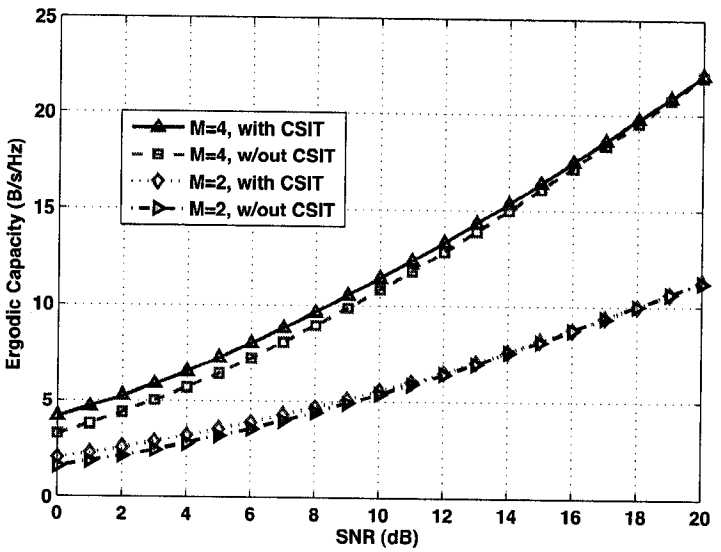

Figure 2.3 Ergodic capacity with and without perfect CSIT as a function of SNR for $M_t = M_r = M = 2, 4$.

2.4 Beamforming or Maximum Ratio Transmission

MIMO systems can be configured to obtain diversity and array gains instead of capacity gain. This type of configuration for multiple-antenna systems is referred to as MIMO *beamforming* or *maximum ratio transmission.* In this configuration, contrary to spatial multiplexing systems (where multiple symbols are simultaneously transmitted), a single symbol, weighted by a complex vector, is sent over each transmit antenna, so that the input covariance matrix has unit rank. An essential requirements of such configuration is the availability of (partial or perfect) channel knowledge at transmitter (CSIT). The term "beamforming" in MIMO communications is different from the similar term in smart antennas and those two must not be mistaken. Literature on point-to-point MIMO beamforming is very rich and a short discussion on this topic deserves separate treatments. But for a quick review, we will mention some of the pioneering works. MIMO beamforming was first considered in [27] in terms of array and diversity gains. In [28], Dighe *et al* analyzed the performance of MIMO beamforming in terms of symbol error probability over Rayleigh channels and in [29] investigated the outage probability over Rayleigh and Rician channels. [30] and [31] have also analyzed the array gain in MIMO channels. All these works have assumed a perfect channel knowledge at the transmitter. But in practice, obtaining the *unlimited-rate, error-free* and *undelayed* channel information at transmitter is challenging (especially for multiple antenna systems). As a result, follow-up works have assumed imperfect, partial or limited channel knowledge at the transmitter and have proposed efficient algorithms to transfer channel information from

receiver to transmitter. Examples of such works are found in [32–37]. In the sequel we present the fundamentals of MIMO beamforming assuming perfect CSIT and perfect CSIR.

2.4.1 System Model

Consider the slow-fading ($M_r \times M_t$) MIMO channel described in section 2.2. To obtain a ($M_t \times 1$) transmit beamforming vector, we need to apply the following transformation:

$$\boldsymbol{w}_t = \frac{1}{a}(\boldsymbol{w}_r \boldsymbol{H})^* \tag{2.36}$$

where $\boldsymbol{w}_r \in \mathbb{C}^{1\times M_r}$ is the receive beamforming vector and $a \in \mathbb{R}$ is a normalization constant defined by:

$$a = |\boldsymbol{w}_r \boldsymbol{H}|. \tag{2.37}$$

We can write the transmit data vector as:

$$\boldsymbol{x} = \boldsymbol{w}_t s \tag{2.38}$$

where s is the transmitted symbol chosen from a complex constellation. The received vector is therefore:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{H}\boldsymbol{x} + \boldsymbol{z} && (2.39)\\ &= \frac{s}{a}\boldsymbol{H}(\boldsymbol{w}_r \boldsymbol{H})^* + \boldsymbol{z} && (2.40) \end{aligned}$$

After processing the received vector by the receive beamforming vector, we obtain the following scalar quantity:

$$\begin{aligned} \tilde{s} &= \frac{s}{a}\boldsymbol{w}_r\boldsymbol{H}(\boldsymbol{w}_r\boldsymbol{H})^* + \boldsymbol{w}_r\boldsymbol{z} && (2.41)\\ &= \frac{s}{a}\boldsymbol{w}_r\boldsymbol{H}\boldsymbol{H}^*\boldsymbol{w}_r^* + \boldsymbol{w}_r\boldsymbol{z} && (2.42)\\ &= as + \boldsymbol{w}_r\boldsymbol{n} && (2.43) \end{aligned}$$

with the signal-to-noise ratio defined by:

$$\mathsf{SNR} = \frac{a^2}{\boldsymbol{w}_r\boldsymbol{w}_r^*}\mathsf{SNR}_0 = a^2\mathsf{SNR}_0, \tag{2.44}$$

where SNR_0 is the average SNR in the case of single transmit and receive antennas and $\boldsymbol{w}_r\boldsymbol{w}_r^* = 1$. From (2.44) we understand that SNR is a function of $\boldsymbol{w}_r$ and maximizing SNR requires optimization over the receive beamforming vector. Consider the following maximization problem

$$\begin{aligned} \arg \quad & \left\{ \max_{\boldsymbol{w}_r} \quad \{\boldsymbol{w}_r\boldsymbol{H}\boldsymbol{H}^*\boldsymbol{w}_r^*\} \right\} && (2.45)\\ & s.t. \qquad |\boldsymbol{w}_r| = 1. \end{aligned}$$

According to Rayleigh-Ritz theorem [38], for any non-zero $N \times 1$ complex vector and a given $N \times N$ Hermitian matrix[2] $\boldsymbol{A}$, $\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x} \leq |\boldsymbol{x}|^2\lambda_{max}$, where λ_{max} is the largest eigenvalue of $\boldsymbol{A}$. The equality holds if and only if $\boldsymbol{x}$ is along the eigenvector corresponding to λ_{max}. Based on this fact, solution to

[2]Matrix $\boldsymbol{A}$ is said to be Hermitian if and only if $\boldsymbol{A} = \boldsymbol{A}^*$.

the above maximization problem is:

$$\boldsymbol{w}_r^* = \boldsymbol{u}_{max} = (u_1, u_2, \ldots, u_{M_r})^T \tag{2.46}$$

where $\boldsymbol{u}_{max}$ is the eigenvector corresponding to the largest eigenvalue of the quadratic form $\boldsymbol{H}\boldsymbol{H}^*$. We thus obtain the following transmit beamforming vectors:

$$\boldsymbol{w}_t = \boldsymbol{v}_{max} = (v_1, v_2, \ldots, v_{M_t})^T \tag{2.47}$$

where $\boldsymbol{v}_{max}$ is the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{H}^*\boldsymbol{H}$. The choice of transmit and receive beamforming vectors gives

$$a^2 = \lambda_{max} \tag{2.48}$$

which indicates that the amount of increase in the SNR is equal to the largest eigenvalue of the matrix $\boldsymbol{H}\boldsymbol{H}^*$. In order to further study the bounds on a^2, it can be shown that [27]:

$$\begin{aligned} a &= \frac{1}{\sqrt{M_r}}\Big(\sum_{i=1}^{M_r}\sum_{j=1}^{M_r}|\sum_{k=1}^{M_t} h_{ik}h_{jk}^*|\Big)^{\frac{1}{2}} \\ &= \frac{1}{\sqrt{M_r}}\Big(\sum_{i=1}^{M_r}\sum_{k=1}^{M_t}|h_{ik}|^2 + \sum_{i=1}^{M_r}\sum_{j=1,i\neq j}^{M_r}|\sum_{k=1}^{M_t} h_{ik}h_{jk}^*|\Big)^{\frac{1}{2}}. \end{aligned} \tag{2.49}$$

2.4.2 Array Gain Analysis

The right-hand term in equation (2.49) is the inner product of the rows of channel matrix. We can consider two extreme cases for that inner product.

First, assume that all the rows are mutually orthogonal. In this case, a becomes

$$a = \frac{1}{\sqrt{M_r}} \Big(\sum_{i=1}^{M_r} \sum_{j=1}^{M_t} |h_{ij}|^2 \Big)^{\frac{1}{2}} \tag{2.50}$$

and

$$\mathbb{E}\{a^2\} = M_t \sigma_h^2, \tag{2.51}$$

where σ_h^2 is the variance of the entries in the channel matrix. In the second case, we assume that all rows of the channel matrix are fully correlated which leads to

$$a = \frac{1}{\sqrt{M_r}} \Big(\sum_{i=1}^{M_r} \sum_{j=1}^{M_r} \sum_{k=1}^{M_t} |h_{jk}|^2 \Big)^{\frac{1}{2}} \tag{2.52}$$

and its expected value (array gain) becomes

$$\mathbb{E}\{a^2\} = M_r M_t \sigma_h^2. \tag{2.53}$$

Now we have found a lower and an upper bound on the average SNR:

$$M_t \sigma_h^2 \mathsf{SNR}_0 \leq \mathbb{E}\{\mathsf{SNR}\} \leq M_t M_r \sigma_h^2 \mathsf{SNR}_0. \tag{2.54}$$

The p.d.f. of SNR and the outage probability $\mathbb{P}_{out}$ for maximum ratio transmission can be derived for Rayleigh [29]:

$$\begin{aligned} p(\mathsf{snr}) \;=\; & \frac{\det\Big[\mathbf{\Psi}_c\Big(\frac{\mathsf{snr}}{\mathsf{SNR}_0\sigma^2}\Big)\Big]}{\sigma^2 \mathsf{SNR}_0 \prod_{k=1}^{s} \Gamma(t-k+1)\Gamma(s-k+1)} \\ & \times Tr\Big(\mathbf{\Psi}_c^{-1}\Big(\frac{\mathsf{snr}}{\mathsf{SNR}_0\sigma^2}\Big)\mathbf{\Psi}_c\Big(\frac{\mathsf{snr}}{\mathsf{SNR}_0\sigma^2}\Big)\Big) U(\mathsf{snr}) \end{aligned} \tag{2.55}$$

$$\mathbb{P}_{out} = \mathbb{P}(\mathsf{SNR} \leq \mathsf{SNR}_{th}) = \frac{\det\left[\mathbf{\Psi}_c\left(\frac{\mathsf{SNR}_{th}}{\mathsf{SNR}_0\sigma^2}\right)\right]}{\prod_{k=1}^{s}\Gamma(t-k+1)\Gamma(s-k+1)} \quad (2.56)$$

where $s = \min(M_t, M_r)$, $t = \max(M_t, M_r)$, $\mathbf{\Psi}_c(x)$ is an $s \times s$ Henkel matrix function of $x \in (0, \infty)$ with entries given by

$$\{\mathbf{\Psi}_c(x)\}_{i,j} = \gamma(t-s+i+j-1, x); \;\; i,j = 1, \ldots, s, \quad (2.57)$$

where $\gamma(.,.)$ is an incomplete gamma function and $U(.)$ is the unit step function.

2.4.3 Diversity Gain Analysis

For a system consisting of $M_t \times M_r$ antennas, the order of diversity is $M_t M_r$; in other words, the probability of error decreases inversely with the $(M_t \times M_r)^{th}$ power of the average SNR. Let us assume a system with BPSK modulation. The channel coefficients are Rayleigh distributed and mutually statistically independent. Referring back to (2.49), its second term on the right hand side is always positive. Therefore, the following inequality holds:

$$a^2 \geq \sum_{i=1}^{M_r}\sum_{j=1}^{M_t}|h_{ij}|^2. \quad (2.58)$$

Thus, the worst error probability $\mathbb{P}_e$ is the one obtained where we have equality in (2.58). To determine $\mathbb{P}_e$, the probability of error conditioned on a set of channel coefficients $\{h_{ij}\}$ must be obtained first. Then the conditional

error probability is averaged over the probability density function of h_{ij}. For Gaussian noise, the conditional error probability is expressed as

$$\mathbb{P}_e(\mathsf{SNR}) = Q(\sqrt{2\mathsf{SNR}}) \tag{2.59}$$

The probability density function $p(\mathsf{snr})$ can be determined from the characteristic function of SNR which turns out to be the characteristic function of χ^2-distributed random variable with $2 \times M_r \times M_t$ degrees of freedom for i.i.d. channel coefficients. It follows that $p(\mathsf{SNR})$ is given by

$$p(\mathsf{snr}) = \frac{\mathsf{snr}^{M_t M_r - 1} \exp(-\frac{\mathsf{snr}}{\mathsf{SNR}_a})}{(M_t.M_r)!\mathsf{SNR}_a^{M_t M_r - 1}} \tag{2.60}$$

where

$$\mathsf{SNR}_a = \mathsf{SNR}_0 \mathbb{E}\{|h_{ij}|^2\} \tag{2.61}$$

$$= \mathsf{SNR}_0 \sigma_h^2 \tag{2.62}$$

The error probability is then given by the following integral:

$$\mathbb{P}_e = \int_0^\infty \mathbb{P}_e(\mathsf{SNR}) p(\mathsf{snr}) d(\mathsf{SNR}). \tag{2.63}$$

For $\mathsf{SNR}_a \gg 1$, we then obtain

$$\mathbb{P}_e \approx \left(\frac{1}{4\mathsf{SNR}_a}\right)^{M_r M_t} \frac{(2M_t M_r - 1)!}{(M_t M_r)!(M_t M_r - 1)!}. \tag{2.64}$$

In Table 2.1, the maximum achievable diversity and array gain in different MIMO configurations are shown. In Figures 2.4 -2.6, BER performance of

uncoded BPSK modulation with maximum ratio transmission is plotted. A Rayleigh channel model is used, where the fading channel coefficients are complex Gaussian, i.e. $\mathcal{H} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I})$. Because the objective of carrying out the simulations is to evaluate the performance, it is assumed that transmitter and receiver are perfectly synchronized and perfect knowledge of channel fading coefficients are available to both. The BER performance curves plotted in these figures show the results of using one and two receiving antennas. Two observations that are particularly associated with diversity is:

1. The improvement becomes greater as SNR increases.
2. The incremental improvement becomes smaller as the diversity order increases.

In Figure 2.6, we have compared the performance of 1×4 MRT system with a 2×2 MRT. Both systems show 4^{th}-order diversity but the 1×4 MRT system exhibits an extra power gain (less than 3 dB) because of the number transmit antennas and exploiting CSIT compared to the 2×2 MRT system.

CONFIGURATIONS	ARRAY GAIN	DIVERSITY ORDER
SIMO (CSIR)	M_r	M_r
SIMO (CSIR, CSIT)	M_r	M_r
MISO (CSIR)	1	M_r
MISO (CSIR, CSIT)	M_t	M_t
MIMO (CSIR)	M_r	$M_r M_t$
MIMO (CSIR, CSIT)	$M_r M_t$	$M_r M_t$

Table 2.1 Array gain and diversity order for different MIMO configurations.

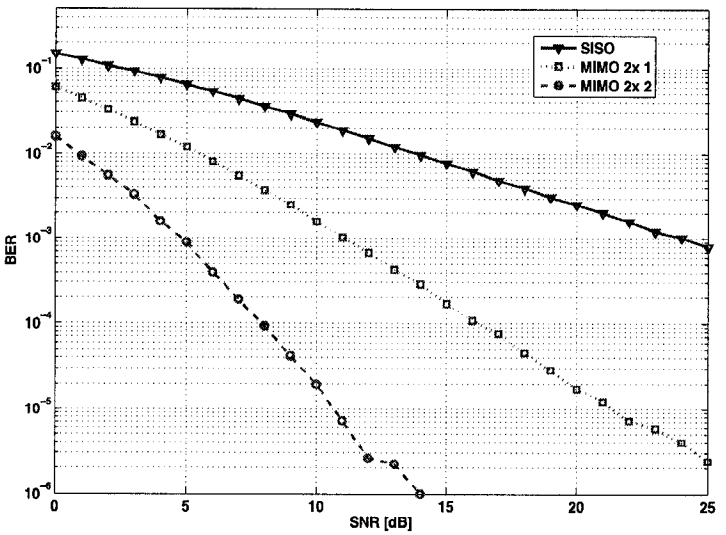

Figure 2.4 BER of uncoded BPSK over Rayleigh channel with different antenna configurations.

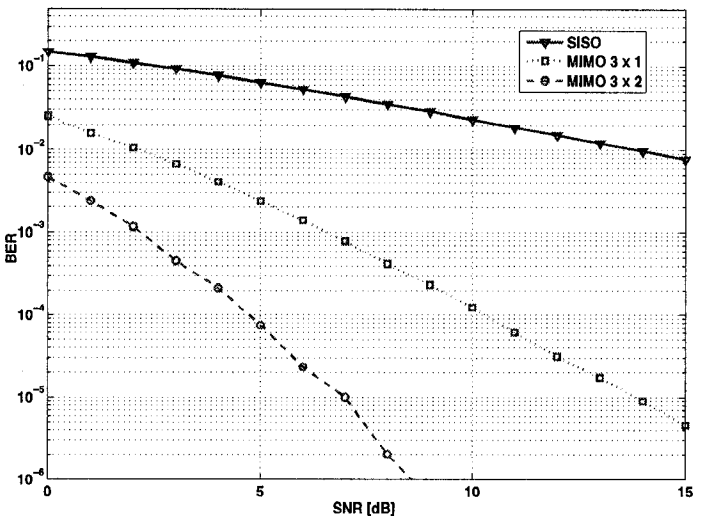

Figure 2.5 BER of uncoded BPSK over Rayleigh channel with different antenna configurations.

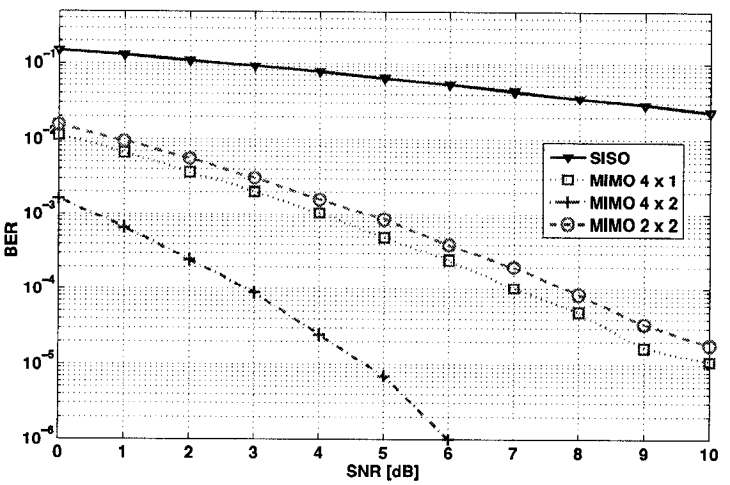

Figure 2.6 BER of uncoded BPSK over Rayleigh channel with different antenna configurations.

CHAPTER 3

INTRODUCTION TO COOPERATIVE COMMUNICATIONS

Despite the indisputable advantages of MIMO communications, there exist certain circumstances under which these benefits can not be completely exploited. Due to the size or cost constraints, not all the mobile nodes are able to be equipped with multiple antennas. Even if there were no such constraints, as we observed in the previous chapter, in highly-correlated line-of-sight MIMO links, capacity gains promised in [4] and [39] would not be achievable. These limitations have motivated researchers to come up with new wireless communications techniques in order to combat these constraints as well as preserving the promised MIMO gains. This new technique is known as *user cooperation* or *cooperative diversity* in the literature [40].

Cooperative diversity is an attractive approach to improve performance of wireless systems by creating distributed virtual antennas across different nodes. In this setup, the basic idea is to leverage the antennas available at other nodes in the network as a source of virtual spatial diversity. The term "*user cooperation diversity* " was first introduced by Sendonaris, Erkip, and Aazhang in [6,7] but the principles behind user cooperation can be traced back to *relay* channels.

The relay channel is a channel in which there is one *source* node, one *destination* node and several *relay* nodes which help the source to better transmit

its data to the destination. In the simplest case, there is only one relay node to improve transmission from source to destination. Relay channel is a combination of a *broadcast channel (BC)*($\mathcal{S} \rightarrow \mathcal{D}, \mathcal{R}$) and a *multiple-access channel (MAC)* ($\mathcal{R}, \mathcal{S} \rightarrow \mathcal{D}$) [24]. Figure 3.1 shows the block diagram of such a channel. The Gaussian relay channel was first introduced by Van der Mullen in [41–43]. Substantial advances in the theory were made by Cover and El Gamal [44], who developed two fundamental coding strategies for relay channels. It was not until recently that the relay channel theory re-gained considerable attention because of its application in *cooperative communication* [6,7,45–47]. Despite recent works which have addressed capacity approaching strategies for single-antenna [48,49] and multiple-antenna relay channels [9], the capacity achieving strategy for a general relay channel is still *unknown*.

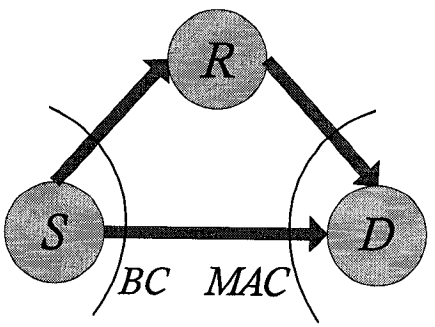

Figure 3.1 Block diagram of a relay channel.

However, there are several differences in the theory of cooperative communication and relay channels. First, in cooperative communication, each node

may act as both a source and a relay (for another node), while in relay channels, the relays are dedicated to help the source and do not have any information of their own to transmit. Second, in relay channels, the main objective is to approach the capacity limits of a Gaussian channel, while in cooperative communication, the main focus is to leverage diversity to overcome channel fading in a distributed manner.

The rest of this chapter is organized as follows: we discuss the capacity bounds of two important classes of relay channels, namely the "full-duplex (FD)" and the "time-division-duplex (TDD)" channels in section 3.1 and section 3.2 respectively. In section 3.3 a variant of time-division-duplex relay channels namely the "half-duplex (HD)" relay channels is scrutinized.

3.1 Full-Duplex (FD) Relay Channels

3.1.1 System Model

A relay node that can transmit and receive simultaneously at the same frequency is said to be in the *full-duplex (FD)* mode. Data transmission in FD relay channels is performed over time frames in which n symbols denoted by $x_{\mathcal{S}}[1], x_{\mathcal{S}}[2], \ldots, x_{\mathcal{S}}[n]$ are broadcasted from the source. Based on the observation of the n received symbols at the relay (denoted by $y_{\mathcal{R}}[1], y_{\mathcal{R}}[2], \ldots, y_{\mathcal{R}}[n]$), n symbols (denoted by $x_{\mathcal{R}}[1], x_{\mathcal{R}}[2], \ldots, x_{\mathcal{R}}[n]$) are transmitted from the relay to the destination. The final input-output relation of a memoryless FD

relay channel may be written by

$$y_{\mathcal{R}}[n] = h_{\mathcal{SR}}x_{\mathcal{S}}[n] + z_{\mathcal{R}}[n] \tag{3.1}$$

$$y_{\mathcal{D}}[n] = h_{\mathcal{SD}}x_{\mathcal{S}}[n] + h_{\mathcal{RD}}x_{\mathcal{R}}[n] + z_{\mathcal{D}}[n] \tag{3.2}$$

where $y_{\mathcal{D}}[n]$ denotes the n^{th} received symbol at destination; $h_{\mathcal{SR}}$, $h_{\mathcal{SD}}$ and $h_{\mathcal{RD}}$ denote the channel gains corresponding respectively to $\mathcal{S} \rightarrow \mathcal{R}$, $\mathcal{S} \rightarrow \mathcal{D}$ and $\mathcal{R} \rightarrow \mathcal{D}$ links; and finally $z_{\mathcal{R}}[n]$ and $z_{\mathcal{D}}[n]$ are respectively i.i.d. unit variance additive white Gaussian noise at relay and destination. There is an average power constraint on each frame for both source and relay described by

$$\frac{1}{n}\sum_{i=1}^{n} x_{\mathcal{S}}[i]^2 = P_{\mathcal{S}} \tag{3.3}$$

$$\frac{1}{n}\sum_{i=1}^{n} x_{\mathcal{R}}[i]^2 = P_{\mathcal{R}}. \tag{3.4}$$

3.1.2 Capacity Bounds

By applying the max-flow min-cut theorem [24], the capacity of a fixed-gain Gaussian relay channel is upper bounded by [44, 48]:

$$C_{FD}^{upper} = \max_{p(x_{\mathcal{S}},x_{\mathcal{R}})} \min\left\{I(x_{\mathcal{S}}; y_{\mathcal{R}}, y_{\mathcal{D}}|x_{\mathcal{R}}), I(x_{\mathcal{S}}, x_{\mathcal{R}}; y_{\mathcal{D}})\right\} \tag{3.5}$$

$$= \max_{0\leq\beta\leq 1} \min\left\{\frac{1}{2}\log\left(1+(1-\beta)P_{\mathcal{S}}(h_{\mathcal{SR}}^2 + h_{\mathcal{SD}}^2)\right), \right.$$
$$\left. \frac{1}{2}\log\left(1 + h_{SD}^2 P_{\mathcal{S}} h_{\mathcal{RD}}^2 P_{\mathcal{R}} + 2\sqrt{\beta h_{\mathcal{SD}}^2 h_{\mathcal{RD}}^2 P_{\mathcal{S}} P_{\mathcal{R}}}\right)\right\} \tag{3.6}$$

A lower bound can be derived for the capacity of Gaussian relay channels with *decode-and-forward (DF)* strategy at the relay. By DF we mean that the relay node fully decodes the received signal from the source node, encodes it using an independent codebook and forwards the encoded data to the destination node. The DF lower bound is given by [48]:

$$C_{FD}^{DF} = \max\{R_1, R_2\}, \tag{3.7}$$

where R_1 and R_2 are defined by

$$\begin{aligned} R_1 &= \max_{p(x_S, x_R)} \min\Big\{I(x_S; y_R|x_R), I(x_S; x_R, y_D)\Big\} && (3.8)\\ &= \max_{0\le\beta\le 1} \min\Big\{\frac{1}{2}\log\Big(1+(1-\beta)P_S(h_{SR}^2)\Big), \\ &\quad \frac{1}{2}\log\Big(1+h_{SD}^2 P_S h_{RD}^2 P_R + 2\sqrt{\beta h_{SD}^2 h_{RD}^2 P_S P_R}\Big)\Big\} && (3.9)\end{aligned}$$

and

$$R_2 = \frac{1}{2}\log\Big(1+h_{SD}^2 P_S + \frac{h_{SR}^2 P_S}{1+\frac{h_{SD}^2 P_S + h_{SR}^2 P_S + 1}{h_{RD}^2 P_R}}\Big). \tag{3.10}$$

3.2 Time Division Duplex (TDD) Relay Channels

3.2.1 System Model

It is very unrealistic in practice for small-sized relay terminals to receive and transmit simultaneously at the same frequency due to hardware implementations. Hence, relays must be assigned orthogonal resources for their

transmit and receive paths. We assume that such separation is performed in the *time* domain; another approach would be to use frequency relaying where the relays receive and transmit in different *frequency* bands. Consider time-division relaying as an example. For a given time interval T_0, the relay captures data from the source for a duration of αT_0 and relays those data in the remaining fraction of time $(1 - \alpha T_0)$ to the destination. This holds in the same way for frequency-division relays; there, the available bandwidth W_0 instead of time is divided into two fractions: αW_0 for capturing the data data and $(1 - \alpha)W_0$ for relaying the data. Next, we discuss the bounds on the capacity of Gaussian TDD relay channels.

3.2.2 Capacity Bounds

Similarly to what was mentioned in subsection 3.1.2, an upper bound was derived on the capacity of a Gaussian TDD relay channel in [48] as:

$$C_{TDD}^{upper} = \max_{0 \leq \beta \leq 1} \min\{R_3(\beta), R_4(\beta)\} \tag{3.11}$$

where $R_3(\beta)$ and $R_4(\beta)$ are defined by:

$$\begin{aligned} R_3(\beta) &= \frac{\alpha}{2} \log\left(1 + (h_{\mathcal{SD}}^2 + h_{\mathcal{RD}}^2) P_{\mathcal{S},1}\right) + \frac{1-\alpha}{2} \log\left(1 + (1-\beta) P_{\mathcal{S},2} h_{SD}^2\right), \\ R_4(\beta) &= \frac{\alpha}{2} \log\left(1 + h_{\mathcal{SD}}^2 P_{\mathcal{S},1}\right) + \\ &\quad \frac{1-\alpha}{2} \log\left(1 + h_{SD}^2 P_{\mathcal{S},2} + h_{\mathcal{RD}}^2 P_{\mathcal{R}} + 2\sqrt{\beta h_{\mathcal{SD}}^2 h_{\mathcal{RD}}^2 P_{\mathcal{S},2} P_{\mathcal{R}}}\right). \end{aligned} \tag{3.12}$$

Furthermore, two lower bounds can be achieved by decode-and-forward (DF) strategy and *compress-and-forward (CF)* strategies at the relay. The DF lower bound may be written as [48]:

$$C_{TDD}^{DF} = \max_{0\leq\beta\leq 1} \min\{R_5(\beta), R_5(\beta)\} \tag{3.13}$$

where R_5 and R_6 are defined by

$$\begin{aligned} R_5(\beta) &= \frac{\alpha}{2}\log(1+h_{\mathcal{SR}}^2 P_{\mathcal{S},1}) && (3.14)\\ &+ \frac{1-\alpha}{2}\log\Big(1+(1-\beta)h_{\mathcal{SD}}^2 P_{\mathcal{S},2}\Big) && (3.15)\end{aligned}$$

and

$$\begin{aligned} R_6(\beta) &= \frac{\alpha}{2}\log\Big(1+h_{\mathcal{SD}}^2 P_{\mathcal{S},1}\Big) \\ &+ \frac{1-\alpha}{2}\Big(\log(1+h_{\mathcal{SD}}^2 P_{\mathcal{S},2}+h_{\mathcal{RD}}^2 P_{\mathcal{R}}+2\sqrt{\beta h_{\mathcal{SD}}^2 h_{\mathcal{RD}}^2 P_{\mathcal{S},2}P_{\mathcal{R}}}\Big) && (3.16)\end{aligned}$$

Instead of DF strategy at the relay node, we may apply *Wyner Ziv* [50] lossy source coding at the relay terminal. This type of strategy is known in the literature as *compress-and-forward* (CF) relaying. The relay node forwards a compressed version of its channel output to the destination node via an error free channel encoding [24, 45]. This information is used as side information at the destination for decoding. The capacity of CF relaying is given by

$$C_{TDD}^{CF} = \frac{\alpha}{2}\log\Big(1+h_{\mathcal{SD}}^2 P_{\mathcal{S},1}+\frac{h_{\mathcal{SR}}^2 P_{\mathcal{S},1}}{1+\sigma_w^2}\Big)+\Big(\frac{1-\alpha}{2}\Big)\log(1+h_{\mathcal{SD}}P_{\mathcal{S},2}) \tag{3.17}$$

where σ_w^2 is the compression noise given by

$$\sigma_w^2 = \frac{1 + h_{\mathcal{SR}}^2 P_{\mathcal{S},1} + h_{\mathcal{SD}}^2 P_{\mathcal{S},1}}{\left((1 + \frac{h_{\mathcal{RD}}^2 P_{\mathcal{R}}}{1 + h_{\mathcal{SD}}^2 P_{\mathcal{S},2}})^{\frac{1-\alpha}{\alpha}} - 1\right)(h_{\mathcal{SD}}^2 P_{\mathcal{S},1} + 1)}. \tag{3.18}$$

3.2.3 Numerical Results

Capacity bounds for FD and TDD relay channels for fixed channel coefficients $h_{\mathcal{RD}} = 0, 10\ dB$, $h_{\mathcal{SD}} = 0\ dB$ is depicted in Figure 3.2 and Figure 3.3 as a function of source to relay channel gains $h_{\mathcal{SR}}$ for $P_{\mathcal{S}} = P_{\mathcal{R}} = 5\ dB$ [48]. Noise distribution at each receiver is assumed to be Gaussian with zero mean and unit varicance. The CF method can be used for all channels and always gives a rate gain[1] over direct transmission, but as $h_{\mathcal{SR}}$ becomes larger compared to $h_{\mathcal{SD}}$, the DF rate is eventually larger than the CF rate. For both cases of $h_{\mathcal{RD}} = 0, 10\ dB$, FD relaying schemes outperform both the TDD relaying and the direct transmission schemes, but only for large source-relay channel gains they achieve better than direct link with double power. In other words, FD relaying schemes offer power gain over direct transmission schemes only when there is a well-conditioned channel from source to relay.

[1]Rate gain is the gain between the achievable rate compared to direct transmission.

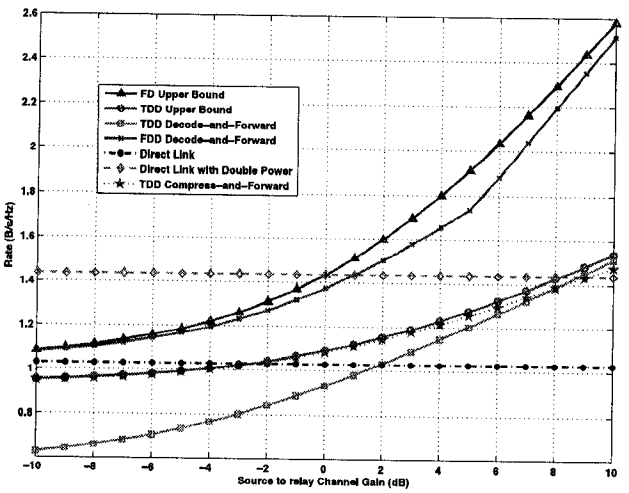

Figure 3.2 Capacity bounds for FD and TDD relay channels with $P_{\mathcal{S}} = P_{\mathcal{R}} = 5dB$, $h_{\mathcal{SD}} = h_{\mathcal{RD}} = 0\ dB$ and $\alpha = 0.4$.

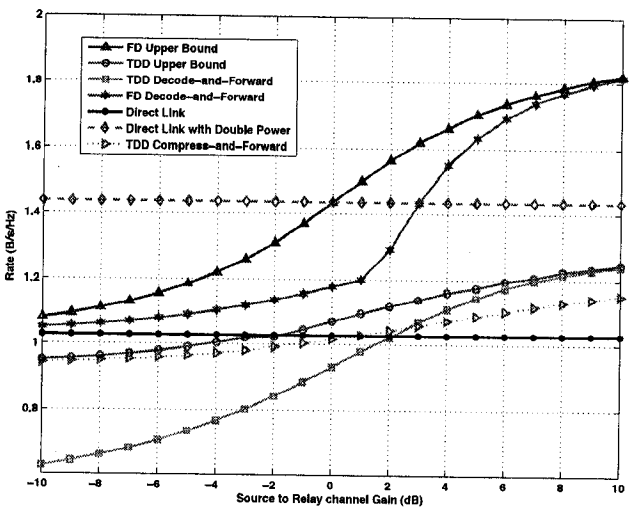

Figure 3.3 Capacity bounds for FD and TDD relay channels with $P_{\mathcal{S}} = P_{\mathcal{R}} = 5dB$, $h_{\mathcal{SD}} = 0\ dB$, $h_{\mathcal{RD}} = 10\ dB$ and $\alpha = 0.4$.

3.3 Half-Duplex Relay Channels

3.3.1 System Model

Among different TDD relaying protocols, half-duplex (HD) relaying (where relays listen during half of the time and transmit over the other half) has gained much more attention in the literature than other TDD relaying schemes. To our best knowledge, pioneering research by Laneman *et al* was the first work on half-duplex relay channels where each relay acts as both information source and a relay for another source [51–53]. The proposed relaying protocols achieve *cooperative diversity* in a distributed manner which compared with *antenna diversity* does not require multiple-antenna nodes. The only drawback of Laneman's protocols is the loss in spectral efficiency, since the source node transmits only during half of the available time or bandwidth.

In [8], two more half-duplex relaying protocols were proposed by Nabar *et al* and analyzed in terms of diversity and multiplexing gains. In [54], these protocols were investigated in terms of diversity-multiplexing tradeoff. In Table 3.1, these three HD relaying protocols are presented. Protocol II is the Laneman's protocol [53] while protocols I and II are the ones proposed by Nabar [8]. As we observe, protocol I is the most inclusive protocol and protocols II and III may be regarded as variants of protocol I. For this reason, in the sequel we study the signal model of protocol I. Extension to protocols II and III is straight-forward.

The source transmits symbol $x_{S,1}$ during the first time slot with power con-

	I	II	III
1	$\mathcal{S} \rightarrow \mathcal{R}, \mathcal{D}$	$\mathcal{S} \rightarrow \mathcal{R}, \mathcal{D}$	$\mathcal{S} \rightarrow \mathcal{R}$
2	$\mathcal{S} \rightarrow \mathcal{D}, \mathcal{R} \rightarrow \mathcal{D}$	$\mathcal{R} \rightarrow \mathcal{D}$	$\mathcal{S} \rightarrow \mathcal{D}, \mathcal{R} \rightarrow \mathcal{D}$

Table 3.1 Three different half-duplex relaying protocols.

straint $\mathbb{E}\{x_{\mathcal{S},1}x^*_{\mathcal{S},1}\} \leq P_{\mathcal{S},1}$. The signals received at the relay and destination are given respectively by

$$y_{\mathcal{R}} = h_{\mathcal{SR}}x_{\mathcal{S},1} + z_{\mathcal{R}} \tag{3.19}$$

and

$$y_{\mathcal{D},1} = h_{\mathcal{SD}}x_{\mathcal{S},1} + z_{\mathcal{D},1} \tag{3.20}$$

where $h_{\mathcal{SR}}$ and $h_{\mathcal{SD}}$ are the channel fading coefficients corresponding to $\mathcal{S} \rightarrow \mathcal{R}$ and $\mathcal{S} \rightarrow \mathcal{D}$ links; and $z_{\mathcal{R}}$ and $z_{\mathcal{D},1}$ are i.i.d. zero-mean unit variance Gaussian noise at source and destination respectively. In the second time slot, the source and relay transmit symbols $x_{\mathcal{S},2}$ and $x_{\mathcal{R}}$ to the destination with power constraints

$$\mathbb{E}\{x_{\mathcal{S},2}x^*_{\mathcal{S},2}\} \leq P_{\mathcal{S},2} \tag{3.21}$$

$$\mathbb{E}\{x_{\mathcal{R}}x^*_{\mathcal{R}}\} \leq P_{\mathcal{R}}. \tag{3.22}$$

The received signal at destination during this time slot is given by

$$y_{\mathcal{D},2} = h_{\mathcal{RD}}x_{\mathcal{R}} + h_{\mathcal{SD}}x_{\mathcal{S},2} + z_{\mathcal{D},2} \tag{3.23}$$

where $h_{\mathcal{RD}}$ and $z_{\mathcal{D},2}$ denote channel fading coefficient for $\mathcal{R} \rightarrow \mathcal{D}$ links and

zero-mean unit-variance Gaussian noise at destination respectively. In what follows, we present the ergodic capacity analysis of these three protocols with amplify-and-forward (AF) and decode-and-forward (DF) relays.

3.3.2 Amplify-and-Forward (AF) Relaying

The AF relaying scheme provides cooperative diversity while minimizing the computational load at the relay node. In AF relaying, the received signal from the source node is simply amplified according to the power constraint at the relay node. Two of the obvious hinderances of the AF relaying scheme are the storage limitations for the large amounts of analog data required for relaying and for combining at the destination, and the noise amplification at the relay node along with the received source signal.

3.3.2.1 Capacity formulation

For the Gaussian HD relay channel the input-output relation can be written as [55]:

$$\underbrace{\begin{bmatrix} y_{\mathcal{D},1} \\ y_{\mathcal{D},2} \end{bmatrix}}_{\boldsymbol{y}_{\mathcal{D}}} = \underbrace{\begin{bmatrix} h_{\mathcal{SD}} & 0 \\ g h_{\mathcal{SR}} h_{\mathcal{RD}} & h_{\mathcal{SD}} \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} x_{\mathcal{S},1} \\ x_{\mathcal{S},2} \end{bmatrix}}_{\boldsymbol{x}_{\mathcal{S}}} + \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & g h_{\mathcal{RD}} & 1 \end{bmatrix}}_{\boldsymbol{B}} \underbrace{\begin{bmatrix} z_{\mathcal{D},1} \\ z_{\mathcal{R}} \\ z_{\mathcal{D},2} \end{bmatrix}}_{\boldsymbol{z}} \tag{3.24}$$

$$\boldsymbol{y}_{\mathcal{D}} = \boldsymbol{A}\boldsymbol{x}_{\mathcal{S}} + \boldsymbol{B}\boldsymbol{z} \tag{3.25}$$

where g is the amplification gain at the relay node. Equation (3.25) is equivalent to a 2×2 MIMO channel with colored noise. The mutual information of a MIMO channel with colored noise is given by [56]:

$$\begin{aligned} I(\boldsymbol{x}_{\mathcal{S}}; \boldsymbol{y}_{\mathcal{D}}) &= h(\boldsymbol{y}_{\mathcal{D}}) - h(\boldsymbol{y}_{\mathcal{D}}|\boldsymbol{x}_{\mathcal{S}}) && (3.26) \\ &= h(\boldsymbol{y}_{\mathcal{D}}) - (h(\boldsymbol{A}\boldsymbol{x}_{\mathcal{S}}|\boldsymbol{x}_{\mathcal{S}}) + h(\boldsymbol{B}\boldsymbol{z}|\boldsymbol{x}_{\mathcal{S}})) && (3.27) \\ &= h(\boldsymbol{y}_{\mathcal{D}}) - h(\boldsymbol{B}\boldsymbol{z}). && (3.28) \end{aligned}$$

From information theory, we know that the entropy of a complex Gaussian random vector is given by [24]:

$$h(\boldsymbol{y}_{\mathcal{D}}) = \log\left((\pi e)^{M_{\mathcal{D}}} \det(\boldsymbol{R}_{\boldsymbol{y}_{\mathcal{D}}})\right) \tag{3.29}$$

where $\boldsymbol{R}_{\boldsymbol{y}_{\mathcal{D}}}$ denotes the covariance matrix of the vector $\boldsymbol{y}_{\mathcal{D}}$:

$$\boldsymbol{R}_{\boldsymbol{y}_{\mathcal{D}}} = \mathbb{E}\{\boldsymbol{y}_{\mathcal{D}}\boldsymbol{y}_{\mathcal{D}}^*\}. \tag{3.30}$$

Here, $\boldsymbol{R}_{\boldsymbol{y}_{\mathcal{D}}}$ can be expressed as:

$$\begin{aligned} \boldsymbol{R}_{\boldsymbol{y}_{\mathcal{D}}} &= \mathbb{E}\{(\boldsymbol{A}\boldsymbol{x}_{\mathcal{S}} + \boldsymbol{B}\boldsymbol{z})(\boldsymbol{A}\boldsymbol{x}_{\mathcal{S}} + \boldsymbol{B}\boldsymbol{z})^*\} && (3.31) \\ &= \mathbb{E}\{\boldsymbol{A}\boldsymbol{x}_{\mathcal{S}}\boldsymbol{x}_{\mathcal{S}}^*\boldsymbol{A}^*\} + \mathbb{E}\{\boldsymbol{B}\boldsymbol{z}\boldsymbol{z}^*\boldsymbol{B}^*\} && (3.32) \\ &= \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}_{\mathcal{S}}}\boldsymbol{A}^* + \boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^* && (3.33) \end{aligned}$$

where $\boldsymbol{R}_{\boldsymbol{x}_S}$ and $\boldsymbol{R}_{\boldsymbol{z}}$ are the covariance matrices of the transmited signal and noise vectors. Substituting $\boldsymbol{R}_{\boldsymbol{y}_D}$ into equation (3.29) yields

$$h(\boldsymbol{y}_D) = \log\Big((\pi e)^{M_D} \det[\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}_S}\boldsymbol{A}^* + \boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^*]\Big) \tag{3.34}$$

Similarly, the entropy of $\boldsymbol{B}\boldsymbol{z}$ becomes

$$h(\boldsymbol{B}\boldsymbol{z}) = \log\Big((\pi e)^{M_D} \det[\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^*]\Big). \tag{3.35}$$

Combining (3.34) and (3.35) into equation (3.26), we obtain the mutual information of the system

$$I(\boldsymbol{x}_S; \boldsymbol{y}_D) = \log\Big(\frac{\det[\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}_S}\boldsymbol{A}^* + \boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^*]}{\det[\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^*]}\Big). \tag{3.36}$$

Exploiting the properties of matrix determinants, a simple expression is obtained

$$I(\boldsymbol{x}_S; \boldsymbol{y}_D) = \log\Big(\det[\boldsymbol{I}_2 + (\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}_S}\boldsymbol{A}^*)(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^*)^{-1}]\Big) \tag{3.37}$$

The capacity of the system is the maximum mutual information between the source and destination subject to the power constraints at source and relay (equations (3.21) and (3.22)). Due to the half-duplex strategy (listen-and-transmit), we are using half of the degrees of freedom of the channel, therefore a factor 1/2 penalty is required:

$$C_{\mathrm{AF}} = \frac{1}{2}\log\Big(\det[\boldsymbol{I}_2 + (\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}_S}\boldsymbol{A}^*)(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{z}}\boldsymbol{B}^*)^{-1}]\Big) \tag{3.38}$$

$$= \frac{1}{2}\log\Big(\det[\boldsymbol{I}_2 + (\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{x}_S}\boldsymbol{A}^*)(\boldsymbol{B}\boldsymbol{B}^*)^{-1}]\Big) \tag{3.39}$$

3.3.2.2 Multiplexing Gain and Diversity Analysis

Assuming equal power allocation over two time slots and $\mathcal{R} \rightarrow \mathcal{D}$ to be AWGN, we may order the information rates supported by the three mentioned protocol for AF relay channels as follows [8]:

$$C_{\mathrm{AF}}^{\mathrm{I}} \geq C_{\mathrm{AF}}^{\mathrm{II}} \geq C_{\mathrm{AF}}^{\mathrm{III}}. \tag{3.40}$$

We shall interpret the above ordering in terms of traditional MIMO gains. From (3.38), we can see that the price to be paid for cooperative transmission over two time slots is a reduction in spectral efficiency (compared with a MIMO system with colocated antennas) accounted for by the factor in front of the log term. Protocol I is the only protocol that can realize a multiplexing gain in the classical sense and, hence, recover (to a certain extent) from this 50% loss in spectral efficiency. We note, however, that the effective channel is not i.i.d. complex Gaussian as is the case in traditional MIMO systems. This implies that in general we may not recover fully from the loss in spectral efficiency. The corresponding difference in performance can be attributed to the fact that we are dealing with a distributed system where the individual terminals have to cooperate through noisy links [57]. Protocols II and III do not provide multiplexing gain, which explains their inferior performance when compared with protocol I. Finally, the fact that protocol II is superior to protocol III can be attributed to the fact that protocol II corresponds to a SIMO system realizing array gain, whereas protocol III corresponds to a MISO system devoid of array gain (recall that we assumed

perfect channel knowledge in the receivers and no channel knowledge in the transmitters). Maximizing the degree of broadcasting at the source and receive collision at the destination (as is done in Protocol I) will in general result in a higher number of degrees-of-freedom (and, hence, higher achievable rates in the degrees-of-freedom limited case) reflected by the creation of an effective MIMO channel [54].

We shall next analyze and compare the different protocols from a diversity point-of-view. Following the approach in [21,53], we shall interpret the outage probability at a certain transmission rate as the packet-error rate (PER). The diversity order is then given by the magnitude of the slope of the PER as a function of SNR (on a log-log scale). To be more precise, we define the diversity order for transmission rate R as:

$$d(R) = \lim_{\mathsf{SNR}\to\infty} \frac{-\mathbb{P}_e(R, \mathsf{SNR})}{\log \mathsf{SNR}}, \tag{3.41}$$

where $\mathbb{P}_e(R, \mathsf{SNR})$ denotes the PER or outage probability at transmission rate R as a function of SNR. Equivalently, a scheme achieving diversity order $d(R)$ at rate R has an error probability $\mathbb{P}_e(R, \mathsf{SNR})$ that behaves as $\mathbb{P}_e(R, \mathsf{SNR}) \propto \mathsf{SNR}^{-d(R)}$ at high SNR. We summarize the results of [8] by noting that

$$\mathrm{PEP}_{\mathrm{AF}}^{\mathrm{I}} \leq \mathrm{PEP}_{\mathrm{AF}}^{\mathrm{II}} \leq \mathrm{PEP}_{\mathrm{AF}}^{\mathrm{III}}, \tag{3.42}$$

and also the fact that all three protocols achieve second-order diversity according to the definition in (3.41). Recall that in traditional MIMO systems

the presence of array gain is reflected by an increased receive SNR when compared with the case where no array gain is present. Consequently, the ordering in (3.42) can be interpreted as reflecting the amount of array gain realized by the individual protocols in the AF mode.

3.3.3 Decode-and-Forward (DF) Relaying

DF type protocols have the advantage that only digital data has to be stored, but on the other hand they require the relay node to fully decode the signal stream which may impose a heavy computational load on the relay when the current prevailing FEC codes (e.g. convolutional, turbo, LDPC codes) are used.

3.3.3.1 Capacity Formulation

The overall input-output relation of a DF relay channel is given by:

$$\underbrace{\begin{bmatrix} y_{\mathcal{D},1} \\ y_{\mathcal{D},2} \end{bmatrix}}_{\boldsymbol{y}_{\mathcal{D}}} = \underbrace{\begin{bmatrix} h_{\mathcal{SD}} & 0 \\ h_{\mathcal{RD}} & h_{\mathcal{SR}} \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} x_{\mathcal{S},1} \\ x_{\mathcal{S},2} \end{bmatrix}}_{\boldsymbol{x}_{\mathcal{S}}} + \underbrace{\begin{bmatrix} z_{\mathcal{D},1} \\ z_{\mathcal{D},2} \end{bmatrix}}_{\boldsymbol{z}} \tag{3.43}$$

Note that equation (3.43) assumes that the relay was able to correctly decode the data received from the source. From network information theory, one can realize that the achievable information rate of DF relaying is determined by the max-flow min-cut theorem. The cut set around the source

forms a broadcast channel and the cut set around the destination forms a parallel channel (due to the orthogonality in time). The system capacity is the minimum value of the capacities of the two cut sets. From [10] we have the following constraints on the achievable rates of a DF relay channel under relaying protocol I:

$$\begin{aligned} R_1 &\leq \log\left(1+P_{\mathcal{S},1}|h_{\mathcal{SR}}|^2\right) = R_{relay}^{max} && (3.44)\\ R_1 &\leq \log\left(1+P_{\mathcal{S},1}|h_{\mathcal{SD}}|^2+P_{\mathcal{R}}|h_{\mathcal{RD}}|^2\right) = R_1^{max} && (3.45)\\ R_2 &\leq \log\left(1+P_{\mathcal{S},2}|h_{\mathcal{SD}}|^2\right) = R_2^{max} && (3.46)\\ R_1+R_2 &\leq \log\det\left(\boldsymbol{I}_2+P_{\mathcal{S}}\boldsymbol{A}\boldsymbol{A}^*\right) = R_{1+2}^{max} && (3.47) \end{aligned}$$

where R_1, R_2 and R_{1+2} are respectively the achievable rates in the first, second and both time slots. The capacity of the HD-DF relay channel is then given by

$$C_{DF} = \begin{cases} R_{1+2}^{max}, & R_{relay}^{max} \geq R_{1+2}^{max} - R_2^{max} \\ R_{relay}^{max} + R_2^{max}, & otherwise \end{cases} \quad (3.48)$$

3.3.3.2 Multiplexing Gain and Diversity Analysis

The relation between the ergodic and outage capacities of the three relaying protocols is as follows [8]:

$$C_{\mathrm{DF}}^{\mathrm{I}} \geq C_{\mathrm{DF}}^{\mathrm{III}} \geq C_{\mathrm{DF}}^{\mathrm{II}}. \quad (3.49)$$

Again, the superiority of protocol I can be attributed to the fact that it realizes multiplexing gain of two in the classical sense and, hence, recovers from (some of) the loss due to the use of two time slots for transmission. On the other hand, protocol II and III achieve multiplexing gain of one. For diversity gain analysis, similar approach as that of AF relaying can be taken. From [8], we can mention that protocol II extracts only first-order diversity in the DF mode which follows intuitively from the fact that the information rate for Protocol II in the DF mode can never exceed that supported by the $\mathcal{S} \rightarrow \mathcal{R}$ Rayleigh fading channel (denoted by R_{relay}^{max}). Despite weak performance of protocol II in terms of diversity, protocol I and III extract second-order diversity in the DF mode. We note that if the $\mathcal{R} \rightarrow \mathcal{D}$ link is assumed fading and the $\mathcal{S} \rightarrow \mathcal{R}$ link is static with large channel gain, so that the $\mathcal{S} \rightarrow \mathcal{R}$ link is not a bottleneck over the first time-slot, then it is straightforward to show that all three protocols are capable of extracting second-order diversity in the DF mode.

3.3.4 Numerical Results

In Figure 3.4, we have plotted the ergodic capacities of AF relaying for the three different protocols described in Table 3.1 as functions of $\mathsf{SNR}_{\mathcal{S}\rightarrow\mathcal{R}}$ for $\mathsf{SNR}_{\mathcal{S}\rightarrow\mathcal{D}} = \mathsf{SNR}_{\mathcal{R}\rightarrow\mathcal{D}} = 10\ dB$. All channel gains $h_{\mathcal{SR}}$, $h_{\mathcal{SD}}$ and $h_{\mathcal{RD}}$ are assumed to be independently Rayleigh distributed. We can see that at low SNR (when there is large noise amplification at the relay terminal), protocol III is the worst among all protocols. At high SNR (when there is negligible noise amplification), protocol II and III perform quite similarly and are

outperformed by protocol I which benefits from multiplexing gain.

Figure 3.5 illustrates the ergodic capacities of three DF protocols as functions of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$ for $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}} = \mathsf{SNR}_{\mathcal{R}\to\mathcal{D}} = 10\ dB$. We observe that protocol II is severely SNR-limited when the $\mathcal{S} \to \mathcal{R}$ is very poor. At high SNR when $\mathcal{S} \to \mathcal{R}$ is in good shape, protocols II and III perform equally and are outperformed by protocol I due to its the multiplexing gain which recovers some of the $1/2$ factor loss in a TDMA-based transmission.

We have plotted ergodic capacities of AF and DF relaying as functions of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$ for different values of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$ and $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$ in Figure 3.6 (for protocol I) and Figure 3.7 (for protocol II). In these Figures, we see that AF relaying is far better than DF relaying when the relay terminal is closer to the destination than the source. On the other hand, when the relay node approaches the source node (i.e. $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$ increases), the DF capacity starts to grow and will finally slightly outperform AF relaying at very high SNR. In Figure 3.8 and Figure 3.9, we have plotted the ergodic capacities of AF and DF relaying as functions of $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$ for different values of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$ and $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$ for protocol I and II, respectively. For protocol I, as the relay node approaches the destination node, AF performance deteriorates due to noise amplification and DF relaying outperforms AF relaying. On the other hand, for protocol II, AF relaying outperforms DF relaying at all SNRs, since it has higher order diversity. A more detailed analysis of the impact of relay's location on the capacity of cooperative channels has been discussed in [58]. For the rest of this thesis, we use HD-AF relaying under protocol II as the basis for our relaying scheme. The reason behind this choice is first the lack

of capacity achieving codes for protocols I and III and the second is the receiver complexity for the destination node under relaying protocols I and III, and also, the general superiority of AF to DF relaying for protocol II. [2]

[2]However, there are some recent advances on coding for non-orthogonal cooperative channels (protocol I) in the literature [59, 60].

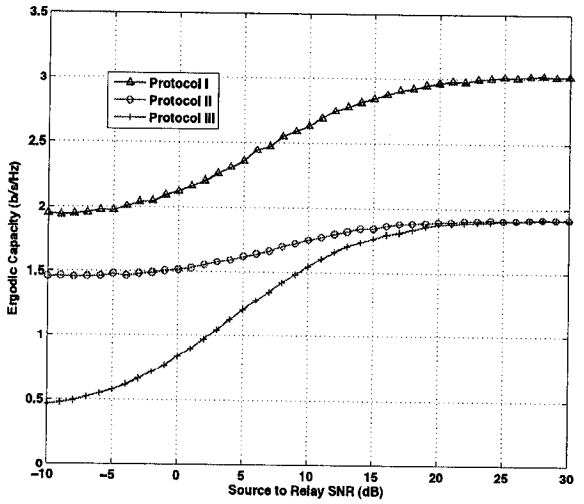

Figure 3.4 Ergodic capacity of different amplify-and-forward (AF) protocols at $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}} = \mathsf{SNR}_{\mathcal{R}\to\mathcal{D}} = 10\ dB$.

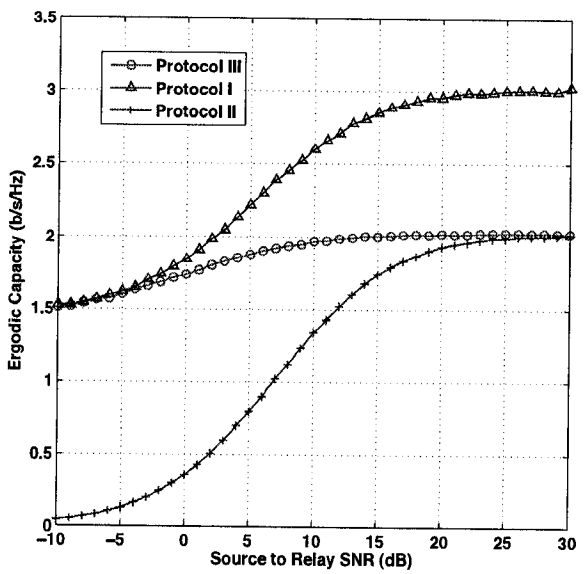

Figure 3.5 Ergodic capacity of different decode-and-forward (DF) protocols at $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}} = \mathsf{SNR}_{\mathcal{R}\to\mathcal{D}} = 10\ dB$.

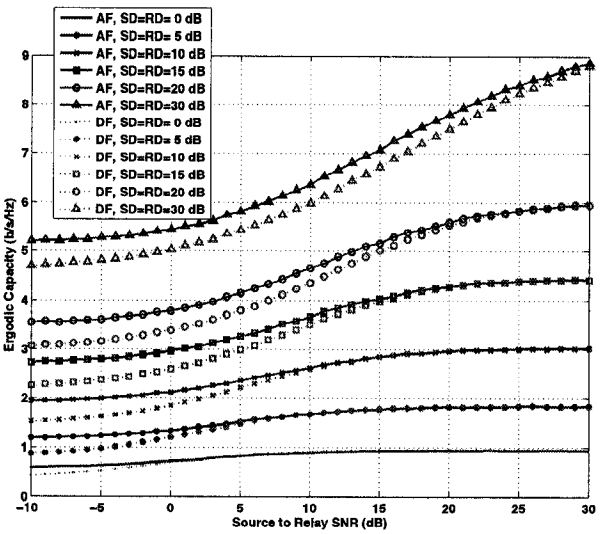

Figure 3.6 Comparison between ergodic capacity of AF and DF relaying operating under protocol I for various values of $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$, $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$ as a function of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$.

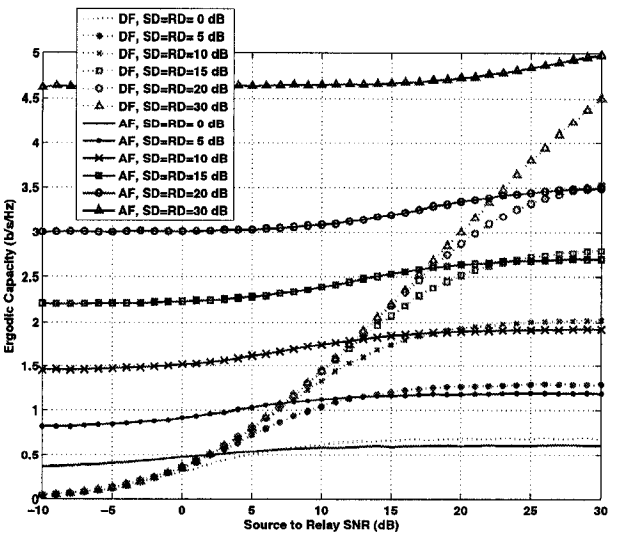

Figure 3.7 Comparison between ergodic capacity of AF and DF relaying operating under protocol II for various values of $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$, $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$ as a function of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$.

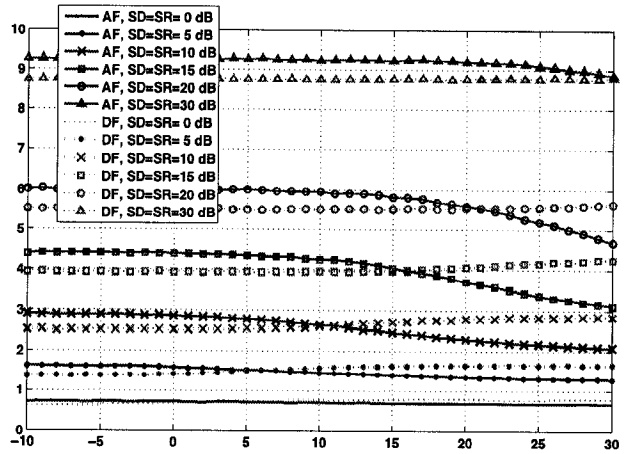

Figure 3.8 Comparison between ergodic capacity of AF and DF relaying operating under protocol I for various values of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$, $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$ as a function of $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$.

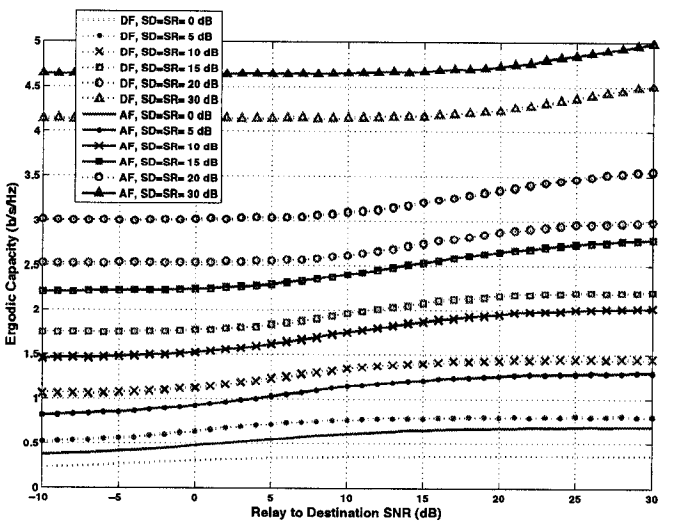

Figure 3.9 Comparison between ergodic capacity of AF and DF relaying operating under protocol II for various values of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$, $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$ as a function of $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$.

CHAPTER 4

MIMO PARALLEL RELAY CHANNELS: RELAY SELECTION AND BEAMFORMING

4.1 Introduction

Cooperative communication has drawn considerable attention from the wireless research community over the last few years [40]. In cooperative communication, multiple relays improve the data transmission from the information source to the destination. Due to the distributed nature of cooperation, a virtual antenna array is formed and, as a result, many of the advantages of multiple-input multiple-output (MIMO) communication, such as gains in throughput, diversity and power, can be achieved. A recent trend in this field is to combine the benefits of MIMO and cooperative communication by employing multiple relays and equipping every node in the network with multiple antennas. This configuration is often referred to as *parallel MIMO relay channel* in the literature [13, 61, 62].

In [10] and [11], a lower bound to the capacity of such a configuration is shown to be achieved through *distributed interference cancellation* in the asymptote of a large number of relays. Follow-up works have reported capacity improvements through different relaying strategies such as backward and forward zero-forcing [63], relaying with phase control [64] and multi-user MMSE re-

laying [65]. Unfortunately, the capacity achieving strategy for MIMO parallel relay channels is still unknown. In a recent work [13], the authors introduced two algorithms namely, *cooperative beamforming scheme (CBS)* and *incremental cooperative beamforming scheme (ICBS)* which achieve the capacity with a small gap in the *asymptotic* case of a large number of relays. However, these algorithms require a large amount of feedback from the relays to the source, need cooperation between the relays and the sum power consumed across the relays is large. A proper approach to resolve such problems is to pre-select a small number of relays by setting a *threshold* and have only relays with large backward and forward channel gains sending their channel information to the source node. Then, a selection criteria is further applied based on *semi-orthogonality* among either *spatial eigenmode* or *antenna* pairs of the pre-selected set of potential relays. Inspired by [66], which shows the benefits of distributed MIMO relaying in the downlink of cellular networks, we apply *zero-forcing multi-user transmit beamforming* at the source to multiplex data streams to the selected subset of relays, and perform *zero-forcing multi-user receive beamfoming* at the destination to recover and separate the received data streams. The simulation results show that the proposed algorithms outperform existing distributed two-hop MIMO relaying schemes for small to medium number of relays. Furthermore, they require lower sum power at the relay nodes and also a smaller amount of feedback in the network. Only as the number of relay nodes goes to infinity, will other algorithms perform better due to the difference in the capacity scaling of point-to-point and point-to-multipoint communications.

4.2 System Model

4.2.1 Network and Protocol Setup

As shown in Figure 4.1, the system under consideration consists of a source $\mathcal{S}$ and a destination $\mathcal{D}$ both equipped with M antennas, and K N-antenna $(K, N \geq M)$ relays $\mathcal{R}_1, \mathcal{R}_2, ..., \mathcal{R}_K$ distributed between the source and the destination. Throughout this chapter we consider *half-duplex* amplify-and-forward (AF) [53] relaying where data transmission from source to destination requires two channel hops and two non-overlapping time slots. During the first time slot, the source transmits several data streams to the relays which are silent. We also assume that the destination does not receive any signal directly from the source because of large-scale fading effects. During the second time slot, while the source node is idle, a selected set of relays amplify and forward their received signal from the source to the destination.

4.2.2 Channel and Signal Model

All channels are considered frequency-flat block fading with independent realization across blocks and the duration of each block length is assumed to be equal to multiple time slots. During the first time slot, $\mathcal{S}$ transmits a multiplexed data vector $\boldsymbol{x}_{\mathcal{S}} \in \mathbb{C}^{M \times 1}$ to *at most* M different relays over the first hop, such that a relay may be assigned with zero, one or even multiple data streams. For presentation simplicity, we write $\mathfrak{R} = \{\mathscr{R}_1, \mathscr{R}_2, \ldots, \mathscr{R}_M\}$ to represent the set of selected relays (note that $\mathscr{R}_i$ and $\mathscr{R}_j$ $(i \neq j)$ may refer

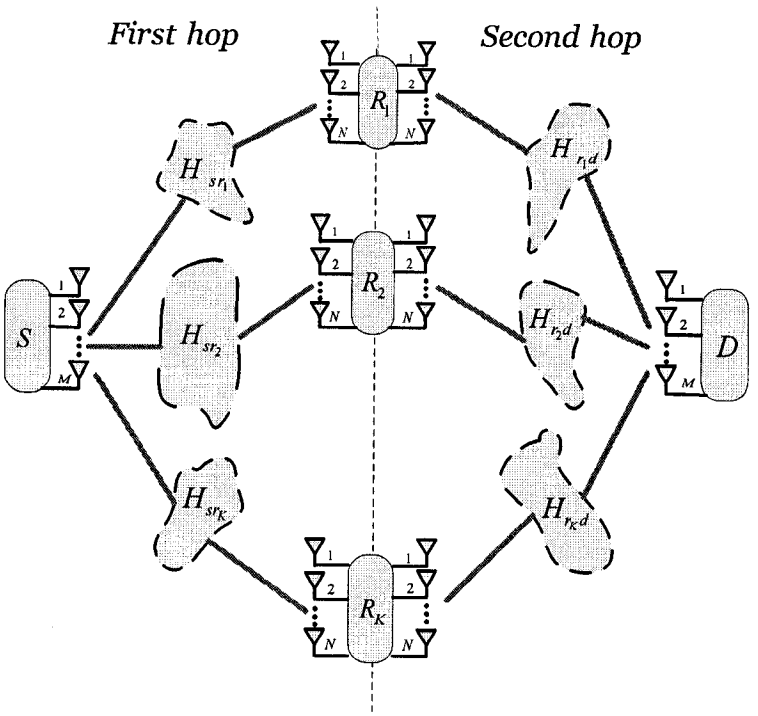

Figure 4.1 Block diagram of a parallel MIMO relay channel.

to the same relay). The source transmit vector can be written as:

$$\boldsymbol{x}_S = \sum_{m=1}^{M} \boldsymbol{w}_{s_m} s_m = \underbrace{[\boldsymbol{w}_{s_1} \boldsymbol{w}_{s_2} \dots \boldsymbol{w}_{s_M}]}_{\boldsymbol{W}_S} \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_M \end{bmatrix}}_{\boldsymbol{s}} \tag{4.1}$$

where s_m is the m^{th} data symbol, $m = 1, 2, \dots, M$ and $\boldsymbol{w}_{s_m} \in \mathbb{C}^{M \times 1}$ is the source transmit beamforming vector corresponding to the m^{th} data symbol. We impose the following long-term power constraint on the source transmit

vector:

$$Tr\Big(\mathbb{E}\{\boldsymbol{x}_{\mathcal{S}}\boldsymbol{x}_{\mathcal{S}}^{*}\}\Big) \leq P_{\mathcal{S}}. \tag{4.2}$$

Suppose that the relay $\mathscr{R}_k$ has been assigned to the k^{th} data stream with the corresponding beamforming vector $\boldsymbol{w}_{s_k}$. We can write the received vector at relay $\mathscr{R}_k$ as:

$$\boldsymbol{y}_{\mathscr{R}_k} = \boldsymbol{H}_{\mathcal{S}\mathscr{R}_k}\boldsymbol{x}_{\mathcal{S}} + \boldsymbol{z}_{\mathscr{R}_k} \tag{4.3}$$

$$= \underbrace{\boldsymbol{H}_{\mathcal{S}\mathscr{R}_k}\boldsymbol{w}_{s_k}s_k}_{Signal} + \underbrace{\boldsymbol{H}_{\mathcal{S}\mathscr{R}_k}\sum_{m=1,m\neq k}^{M}\boldsymbol{w}_{s_m}s_m}_{Interference} + \underbrace{\boldsymbol{z}_{\mathscr{R}_k}}_{Noise} \tag{4.4}$$

where $\boldsymbol{H}_{\mathcal{S}\mathscr{R}_k} \in \mathbb{C}^{N\times M}$ and $\boldsymbol{z}_{\mathscr{R}_k} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N) \in \mathbb{C}^{N\times 1}$ denote the $\mathcal{S} \rightarrow \mathscr{R}_k$ channel matrix and the white zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise vector at relay $\mathscr{R}_k$, respectively. Relay $\mathscr{R}_k$ processes its received vector $\boldsymbol{y}_{\mathscr{R}_k}$ linearly by a matrix $\boldsymbol{G}_{\mathscr{R}_k} \in \mathbb{C}^{N\times N}$ and forwards it to the destination. That i.e. the transmitted vector at relay $\mathscr{R}_k$ is given by:

$$\boldsymbol{x}_{\mathscr{R}_k} = \boldsymbol{G}_{\mathscr{R}_k}\boldsymbol{y}_{\mathscr{R}_k}. \tag{4.5}$$

The aggregate power constraint on the transmit vectors from the relay set $\mathfrak{R}$ is:

$$\sum_{k=1}^{M} Tr\Big(\mathbb{E}\{\boldsymbol{x}_{\mathscr{R}_k}\boldsymbol{x}_{\mathscr{R}_k}^{*}\}\Big) = \sum_{k=1}^{M} Tr\Big(\mathbb{E}\{(\boldsymbol{G}_{\mathscr{R}_k}\boldsymbol{y}_{\mathscr{R}_k})(\boldsymbol{G}_{\mathscr{R}_k}\boldsymbol{y}_{\mathscr{R}_k})^{*}\}\Big) \leq P_{\mathcal{R}}. \tag{4.6}$$

The transmit vectors from the selected relays ($\forall\mathscr{R}_k \in \mathfrak{R}$) then cross the $\mathscr{R}_k \rightarrow \mathcal{D}$ channels (which is be denoted by $\boldsymbol{H}_{\mathscr{R}_k\mathcal{D}} \in \mathbb{C}^{M\times N}$) and add up with

a white ZMCSCG noise vector $\boldsymbol{z}_{\mathcal{D}} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{I}_M) \in \mathbb{C}^{M\times 1}$ at the destination. Assuming perfect synchronization between relays we have the following received vector at destination:

$$\boldsymbol{y}_{\mathcal{D}} = \sum_{k=1}^{M} \boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}\boldsymbol{x}_{\mathcal{R}_k} + \boldsymbol{z}_{\mathcal{D}}. \tag{4.7}$$

The destination multiplies $\boldsymbol{y}_{\mathcal{D}}$ with a receive matrix $\boldsymbol{W}_{\mathcal{D}} \in \mathbb{C}^{M\times M}$ to separate the interfering data streams:

$$\tilde{\boldsymbol{s}} = \boldsymbol{W}_{\mathcal{D}}\boldsymbol{y}_{\mathcal{D}}, \tag{4.8}$$

where $\tilde{\boldsymbol{s}} \in \mathbb{C}^{M\times 1}$ is the vector of estimated data symbols.

4.2.3 Channel State Information (CSI) Model

All relays are assumed to know their local backward and forward channels but not those of other relays. These channel informations can be found for example from *initial* training sequences transmitted from source and destination in a time-division-duplex (TDD) system. After assessing the backward and forward channels, the relays with strong channels will feedback their CSI to the source. The source selects a set of relays (according to the algorithms to be discussed in section 4.3) and then transmits the *second* round of CSI to the relays which consists of information on the processing matrices of the selected relays. Upon reception, the set of selected relays forward those secondary CSI to the destination to build its receive beamforming matrix and

decode the data streams (see Figure 4.2).

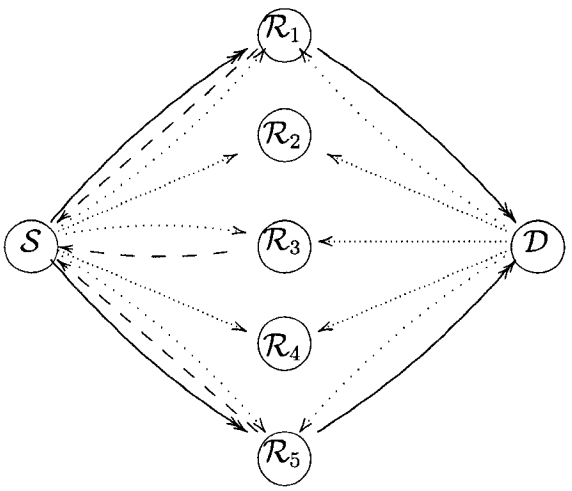

Figure 4.2 A schematics of the CSI model of a TDD system. $\mathcal{S}$ and $\mathcal{D}$ send initial training sequences to the relays ($\cdots >$). $\mathcal{R}_1$, $\mathcal{R}_4$ and $\mathcal{R}_5$ have strong channels and feedback their CSI to $\mathcal{S}$ ($-- >$). From the feedback information, $\mathcal{S}$ selects $\mathcal{R}_1$ and $\mathcal{R}_5$ and sends them the second round of CSI which consists of their processing matrices $\boldsymbol{G}_{\mathcal{R}_k}$ ($— \gg$). $\mathcal{R}_1$ and $\mathcal{R}_5$ will then send those secondary CSI to the destination to build its receive beamforming matrix $\boldsymbol{W}_{\mathcal{D}}$.

4.3 Beamforming and Relay Selection Algorithms

4.3.1 Eigenmode Combining

In this section, we discuss a relay selection algorithm based on selecting a *semi-orthogonal* subset of relays. The idea of selecting a semi-orthogonal subset of nodes in MIMO multi-user network has been investigated extensively in the literature [67], [68]. The rationale behind this idea can be illustrated by a simple example. Consider a single-user MIMO channel as described in chapter 2. The capacity of this channel with *zero forcing* receiver is described by [69]:

$$C_{ZF}(\boldsymbol{\mathcal{H}}) = \sum_{k} \log\left(1 + \frac{\rho_k}{[\boldsymbol{\mathcal{H}}^*\boldsymbol{\mathcal{H}}]^{-1}_{kk}}\right) \tag{4.9}$$

where ρ_k is the water-filling solution, and $\frac{1}{[\boldsymbol{\mathcal{H}}^*\boldsymbol{\mathcal{H}}]^{-1}_{kk}}$ have replaced d_{ii}'s in (2.34). In general, we have $C_{ZF}(\boldsymbol{\mathcal{H}}) \leq C(\boldsymbol{\mathcal{H}})$ since the zero forcing receiver detects the data streams independently. Equality occurs when $\boldsymbol{\mathcal{H}}^*\boldsymbol{\mathcal{H}}/||\boldsymbol{\mathcal{H}}||_F^2 = \boldsymbol{I}$, or equivalently when $\boldsymbol{\mathcal{H}}$ has columns that are orthogonal. This simple example motivates the search for semi-orthogonal channels in MIMO communications.

To achieve a good performance by using zero-forcing beam-forming, the selected sub-channels must have high gains and be nearly orthogonal to each other. As the number of relays increases, it becomes easier to satisfy these requirements. However, the exhaustive search for selecting the best set of users is very complex. Furthermore, the source node must have perfect CSI for backward and forward channels of all relay nodes. To avoid the huge amount of CSI feedback to the source node, we follow the selection algo-

rithm proposed in [70]. To do so, we consider the first hop as a MIMO-BC and the second hop as a MIMO-MAC and propose an efficient sub-optimum algorithm that assigns the coordinates of transmission space to different relays. We will assume that zero-forcing beam-forming is used at the source and destination nodes as the precoding and receiving schemes, respectively. The algorithm starts by setting a threshold value. By applying Singular Value Decomposition (SVD) [38] to backward and forward channels of all relay channel matrices, only the eigenvectors whose corresponding singular values are above the set threshold are considered. Then, among these candidate eigenvectors, the algorithm chooses a set of size M which are nearly orthogonal to each other. As a result of this orthogonality, and the fact that the selected dimensions have large gain values, one can then achieve a good performance in terms of sum-rate throughput. The proposed algorithm is described in details in the following:

Let the singular value decomposition of relay $\mathcal{R}_k$'s backward and forward channel be:

$$\boldsymbol{H}_{\mathcal{S}\mathcal{R}_k} = \boldsymbol{U}_{\mathcal{S}\mathcal{R}_k}\Lambda^{\frac{1}{2}}_{\mathcal{S}\mathcal{R}_k}\boldsymbol{V}^*_{\mathcal{S}\mathcal{R}_k} \tag{4.10}$$

$$\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}} = \boldsymbol{U}_{\mathcal{R}_k\mathcal{D}}\Lambda^{\frac{1}{2}}_{\mathcal{R}_k\mathcal{D}}\boldsymbol{V}^*_{\mathcal{R}_k\mathcal{D}} \tag{4.11}$$

where $\Lambda^{\frac{1}{2}}_{\mathcal{S}\mathcal{R}_k} \in \mathbb{C}^{N\times M}$ and $\Lambda^{\frac{1}{2}}_{\mathcal{R}_k\mathcal{D}} \in \mathbb{C}^{M\times N}$ are diagonal matrices comprising the eigenvalues of $\boldsymbol{H}_{\mathcal{S}\mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}$ respectively, and $\boldsymbol{U}_{\mathcal{S}\mathcal{R}_k} \in \mathbb{C}^{N\times N}$, $\boldsymbol{U}_{\mathcal{R}_k\mathcal{D}} \in \mathbb{C}^{M\times M}$, $\boldsymbol{V}_{\mathcal{S}\mathcal{R}_k} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{V}_{\mathcal{R}_k\mathcal{D}} \in \mathbb{C}^{N\times N}$ are unitary matrices of eigenvectors.

1. *Initialization*: We define a set of *modes* consisting of three indices indicating the index of the relay, the index of the backward eigenmode and the index of the forward eigenmode. The selected set (search domain) is described by:

$$\mathcal{T}_0 = \left\{ (\mathcal{R}_k, i, j) \middle| \lambda_{\mathcal{SR}_k}(i) \geq \lambda_0; \lambda_{\mathcal{R}_k\mathcal{D}}(j) \geq \lambda_0 \right\} \tag{4.12}$$

where $\lambda_{\mathcal{SR}_k}(i)$ and $\lambda_{\mathcal{R}_k\mathcal{D}}(j)$ are the i^{th} and j^{th} singular values of $\boldsymbol{H}_{\mathcal{SR}_k}\boldsymbol{H}^*_{\mathcal{SR}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}\boldsymbol{H}^*_{\mathcal{R}_k\mathcal{D}}$, respectively, and λ_0 is a threshold value. The goal of this threshold is to limit the search space for semi-orthogonal eigenmodes and, most importantly, to constraint the amount of CSI feedback from the relays to the source. The threshold value λ_0 is discussed in section 4.4.

2. *Start-up*: In the second step, after CSI feedback, the source node chooses the mode with the largest effective channel gain. The metric for the effective channel gain of a mode is the ratio of product and sum of backward and forward channel gains of that mode (which has been previously proposed in [71] for single-antenna relays):

$$(\mathscr{R}_1, b_1, f_1) = \arg \max_{(\mathcal{R}_k,i,j)\in\mathcal{T}_0} \frac{\lambda_{\mathcal{SR}_k}(i)\lambda_{\mathcal{R}_k\mathcal{D}}(j)}{\lambda_{\mathcal{SR}_k}(i) + \lambda_{\mathcal{R}_k\mathcal{D}}(j)} \tag{4.13}$$

where $(\mathscr{R}_1, b_1, f_1)$ refers to relay $\mathscr{R}_1$ receiving from its backward channel eigenmode b_1 and transmitting over its forward channel eigenmode f_1. After selecting $(\mathscr{R}_1, b_1, f_1)$, we eliminate this mode from the search

domain $\mathcal{T}_0$ to obtain a new search domain $\mathcal{T}_1$:

$$\mathcal{T}_1 = \mathcal{T}_0 - \{(\mathscr{R}_1, b_1, f_1)\}. \tag{4.14}$$

Besides that for $\forall(\mathcal{R}_k, i, j) \in \mathcal{T}_1$ we define:

$$\Gamma_{(i,j,k)}(1) = \max\left\{\gamma'_{1,k}(i), \gamma''_{1,k}(j)\right\} \tag{4.15}$$

where $\gamma'_{m,k}(i)$ and $\gamma''_{m,k}(j)$ are defined as:

$$\gamma'_{m,k}(i) = \varphi\Big(\boldsymbol{v}_{\mathcal{S}\mathscr{R}_m}(b_m), \boldsymbol{v}_{\mathcal{S}\mathcal{R}_k}(i)\Big) \tag{4.16}$$

$$\gamma''_{m,k}(j) = \varphi\Big(\boldsymbol{u}_{\mathscr{R}_m\mathcal{D}}(f_m), \boldsymbol{u}_{\mathcal{R}_k\mathcal{D}}(j)\Big) \tag{4.17}$$

and $\boldsymbol{v}_{\mathcal{S}\mathcal{R}_k}(i)$ and $\boldsymbol{u}_{\mathcal{R}_k\mathcal{D}}(j)$ are the i^{th} right and j^{th} left eigenvectors of $\boldsymbol{H}_{\mathcal{S}\mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}$, respectively, and $\varphi(\boldsymbol{a}, \boldsymbol{b})$ is a measure of orthogonality between vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ defined as:

$$\varphi(\boldsymbol{a}, \boldsymbol{b}) = \frac{|\boldsymbol{a}^*\boldsymbol{b}|}{|\boldsymbol{a}|.|\boldsymbol{b}|}. \tag{4.18}$$

3. *Iterations*: For $2 \leq m \leq M$, the source performs the following steps:

$$(\mathscr{R}_m, b_m, f_m) = \arg \min_{(\mathcal{R}_k,i,j)\in\mathcal{T}_{m-1}} \Gamma_{(i,j,k)}(m-1) \tag{4.19}$$

$$\mathcal{T}_m = \mathcal{T}_{m-1} - \{(\mathscr{R}_m, b_m, f_m)\} \tag{4.20}$$

$$\Gamma_{(i,j,k)}(m) = \Gamma_{(i,j,k)}(m-1) + \max\left\{\gamma'_{m,k}(i), \gamma''_{m,k}(j)\right\}, \tag{4.21}$$

$$\forall(\mathcal{R}_k, i, j) \in \mathcal{T}_m.$$

Now for the k^{th} selected mode, we may write the processing matrix as:

$$\boldsymbol{G}_{\mathscr{R}_k} = g_{\mathscr{R}_k} \boldsymbol{v}_{\mathscr{R}_k \mathcal{D}}(f_k) \boldsymbol{u}^*_{\mathcal{S}\mathscr{R}_k}(b_k) \tag{4.22}$$

where the coefficients $g_{\mathscr{R}_k}, k = 1, 2, \ldots, M$ satisfy the aggregate power constraint in equation (4.6). Considering the processing matrices $\boldsymbol{G}_{\mathscr{R}_k}$, $k = 1, 2, \ldots, M$, we obtain the following *equivalent* input-output relationship for the first hop:

$$\begin{aligned} r_k &= \boldsymbol{u}^*_{\mathcal{S}\mathscr{R}_k}(b_k) \boldsymbol{H}_{\mathcal{S}\mathscr{R}_k} \boldsymbol{x}_{\mathcal{S}} + \boldsymbol{u}^*_{\mathcal{S}\mathscr{R}_k}(b_k) \boldsymbol{z}_{\mathscr{R}_k} \\ &= \sqrt{\lambda_{\mathcal{S}\mathscr{R}_k}(b_k)} \boldsymbol{v}^*_{\mathcal{S}\mathscr{R}_k}(b_k) \boldsymbol{x}_{\mathcal{S}} + z'_k; \quad k = 1, 2, \ldots, M \end{aligned} \tag{4.23}$$

where r_k is the signal plus interference at $\mathscr{R}_k$, and z'_k is a zero-mean unit-variance AWGN at the output of filter $\boldsymbol{u}^*_{\mathcal{S}\mathscr{R}_k}(b_k)$. Appending the above equations, we can re-write them in vector format:

$$\Longrightarrow \boldsymbol{r} = \boldsymbol{\mathcal{H}}_{\mathcal{S}\mathcal{R}} \boldsymbol{x}_{\mathcal{S}} + \boldsymbol{z}_{\mathcal{R}}, \tag{4.24}$$

where $\boldsymbol{r} = [r_1, \ldots, r_M]^T$, $\boldsymbol{z}_{\mathcal{R}} = [z'_1, \ldots, z'_M]^T$ is a ZMCSCG noise vector, and

$$\boldsymbol{\mathcal{H}}_{\mathcal{S}\mathcal{R}} = \begin{bmatrix} \sqrt{\lambda_{\mathcal{S}\mathscr{R}_1}(b_1)} \boldsymbol{v}^*_{\mathcal{S}\mathscr{R}_1}(b_1) \\ \sqrt{\lambda_{\mathcal{S}\mathscr{R}_2}(b_2)} \boldsymbol{v}^*_{\mathcal{S}\mathscr{R}_2}(b_2) \\ \vdots \\ \sqrt{\lambda_{\mathcal{S}\mathscr{R}_M}(b_M)} \boldsymbol{v}^*_{\mathcal{S}\mathscr{R}_M}(b_M) \end{bmatrix} \in \mathbb{C}^{M \times M}, \tag{4.25}$$

is the equivalent channel matrix of the first hop. Relay $\mathscr{R}_k$ amplifies the received signal r_k by a factor of $g_{\mathscr{R}_k}$to obtain t_k and then filters it through

the filter $\boldsymbol{v}_{\mathscr{R}_k\mathcal{D}}(f_k)$. Re-writing equation (4.7), we obtain the equivalent input-output relationship of the second hop:

$$\begin{aligned}\boldsymbol{y}_{\mathcal{D}} &= \sum_{k=1}^{M} \boldsymbol{H}_{\mathscr{R}_k\mathcal{D}}\boldsymbol{v}^*_{\mathscr{R}_k\mathcal{D}}(f_k)t_k + \boldsymbol{z}_{\mathcal{D}} \\ &= \sum_{k=1}^{M} \sqrt{\lambda_{\mathscr{R}_k\mathcal{D}}(f_k)}\boldsymbol{u}_{\mathscr{R}_k\mathcal{D}}(f_k)t_k + \boldsymbol{z}_{\mathcal{D}}. \end{aligned} \tag{4.26}$$

We may write the above equation in a simpler format:

$$\Longrightarrow \boldsymbol{y}_{\mathcal{D}} = \boldsymbol{\mathcal{H}}_{\mathcal{R}\mathcal{D}}\boldsymbol{t} + \boldsymbol{z}_{\mathcal{D}}, \tag{4.27}$$

where $\boldsymbol{t} = [t_1, \ldots, t_M]^T$ and

$$\boldsymbol{\mathcal{H}}_{\mathcal{R}\mathcal{D}} = \begin{bmatrix} \sqrt{\lambda_{\mathscr{R}_1\mathcal{D}}(f_1)}\boldsymbol{u}^T_{\mathscr{R}_1\mathcal{D}}(f_1) \\ \sqrt{\lambda_{\mathscr{R}_2\mathcal{D}}(f_2)}\boldsymbol{u}^T_{\mathscr{R}_2\mathcal{D}}(f_2) \\ \vdots \\ \sqrt{\lambda_{\mathscr{R}_M\mathcal{D}}(f_M)}\boldsymbol{u}^T_{\mathscr{R}_M\mathcal{D}}(f_M) \end{bmatrix}^T \in \mathbb{C}^{M\times M} \tag{4.28}$$

is the equivalent channel matrix of the second hop.

Since we want the k^{th} mode to only amplify and forward the symbol s_k, we apply a zero-forcing matrix at the source regarding the effective channel matrix of the first hop to pre-cancel the inter-stream interference term. A proper choice for the zero-forcing matrix is the pseudo-inverse of the equivalent channel matrix; which in the case of a square channel matrix equals the

inverse of the matrix [38]:

$$\boldsymbol{W}_{\mathcal{S}} = \boldsymbol{\mathcal{H}}_{\mathcal{SR}}{}^{\dagger} = \boldsymbol{\mathcal{H}}_{\mathcal{SR}}{}^{-1}. \tag{4.29}$$

For the second hop, instead of using successive interference cancellation or MMSE receive techniques, the destination multiplies its received vector $\boldsymbol{y}_{\mathcal{D}}$ by a zero-forcing matrix $\boldsymbol{W}_{\mathcal{D}}$ to separate the data streams and fully cancel the inter-stream interference. Similarly to the first hop, the zero-forcing matrix at destination becomes the pseudo-inverse of the equivalent channel matrix:

$$\boldsymbol{W}_{\mathcal{D}} = \boldsymbol{\mathcal{H}}_{\mathcal{RD}}{}^{\dagger} = \boldsymbol{\mathcal{H}}_{\mathcal{RD}}{}^{-1}. \tag{4.30}$$

Using linear beamforming matrices (i.e. $\boldsymbol{W}_{\mathcal{D}}$ and $\boldsymbol{W}_{\mathcal{S}}$) results in M parallel compound sub-channels from the source to the destination. The effective SNR of the k^{th} compound sub-channel will be [63]:

$$\eta_k = \frac{\eta_{\mathcal{S}\mathscr{R}_k}\eta_{\mathscr{R}_k\mathcal{D}}}{1 + \eta_{\mathcal{S}\mathscr{R}_k} + \eta_{\mathscr{R}_k\mathcal{D}}}, \tag{4.31}$$

where $\eta_{\mathcal{S}\mathscr{R}_k}$ and $\eta_{\mathscr{R}_k\mathcal{D}}$ denote respectively the effective SNR corresponding to the first-hop and the second-hop components of the k^{th} compound sub-channel. Consequently the Shannon capacity of these M parallel compound sub-channels may be written as the following maximization problem:

$$C = \mathbb{E}_{\boldsymbol{\mathcal{H}}_{\mathcal{SR}},\boldsymbol{\mathcal{H}}_{\mathcal{RD}}}\left\{ \max_{\eta_k} \sum_{k=1}^{M} (0.5)\log(1+\eta_k) \right\} \tag{4.32}$$

subject to the power constraints in (4.2) and (4.6). We may subsitute (4.31)

into (4.32) and obtain:

$$C = \mathbb{E}_{\mathcal{H}_{\mathcal{SR}},\mathcal{H}_{\mathcal{RD}}}\left\{ \max_{\eta_{\mathcal{SR}_k},\eta_{\mathcal{R}_k\mathcal{D}}} \sum_{k=1}^{M}(0.5)\log\left(\frac{(1+\eta_{\mathcal{SR}_k})(1+\eta_{\mathcal{R}_k\mathcal{D}})}{1+\eta_{\mathcal{SR}_k}+\eta_{\mathcal{R}_k\mathcal{D}}}\right)\right\}$$
$$s.t. \quad \sum_{k=1}^{M}\zeta_k^{'}\sigma_{s_k}^2 \leq P_{\mathcal{S}}, \quad \sum_{k=1}^{M}\sigma_{\mathscr{R}_k}^2 \leq P_{\mathcal{R}}. \tag{4.33}$$

where $\sigma_{s_k}^2$ and $\sigma_{\mathscr{R}_k}^2$ denote respectively the amount of transmit power allocated to the symbol s_k and to the relay $\mathscr{R}_k$'s output signal t_k:

$$\sigma_{s_k}^2 = \mathbb{E}\{s_k s_k^*\}, \tag{4.34}$$
$$\sigma_{\mathscr{R}_k}^2 = \mathbb{E}\{t_k t_k^*\} \tag{4.35}$$

and $\eta_{\mathcal{SR}_k} = \sigma_{s_k}^2$, $\eta_{\mathscr{R}_k\mathcal{D}} = \frac{\sigma_{\mathscr{R}_k}^2}{\zeta_k^{'}}$, and [72],[69]

$$\zeta_k^{'} = \left[\left(\boldsymbol{\mathcal{H}}_{\mathcal{SR}}^*\boldsymbol{\mathcal{H}}_{\mathcal{SR}}\right)^{-1}\right]_{k,k}, \tag{4.36}$$
$$\zeta_k^{''} = \left[\left(\boldsymbol{\mathcal{H}}_{\mathcal{RD}}^*\boldsymbol{\mathcal{H}}_{\mathcal{RD}}\right)^{-1}\right]_{k,k}. \tag{4.37}$$

In fact, $\zeta_k^{'}$ and $\zeta_k^{''}$ represent respectively, the amount of channel gain reduction at the source and noise amplification at the destination due to zero-forcing. The solution to this non-concave optimization problem requires an iterative approach which has been discussed in [73, 74]. For convenience, we assume *uniform* power allocation across each sub-channel at the source and the relays

which gives:

$$\eta_{\mathcal{SR}_k} = \frac{P_{\mathcal{S}}}{Tr\left[\left(\boldsymbol{\mathcal{H}}^*_{\mathcal{SR}}\boldsymbol{\mathcal{H}}_{\mathcal{SR}}\right)^{-1}\right]}, \tag{4.38}$$

$$\eta_{\mathcal{R}_k\mathcal{D}} = \frac{P_{\mathcal{R}}}{Tr\left[\left(\boldsymbol{\mathcal{H}}^*_{\mathcal{RD}}\boldsymbol{\mathcal{H}}_{\mathcal{RD}}\right)^{-1}\right]}, \tag{4.39}$$

$$g^2_{\mathcal{R}_k} = \frac{\sigma^2_{\mathcal{R}_k}}{1+\eta_{\mathcal{SR}_k}}. \tag{4.40}$$

4.3.2 Antenna Combining

In this section, we consider relay selection based on relays semi-orthogonality a technique called *antenna combining* [75]. All the operations in this algorithm are the same as the previous algorithm with the difference that instead of searching for semi-orthogonality among backward and forward channel eigenmodes of the relays, we search among *rows* and *columns* of backward and forward channels of the relays, respectively. Indeed, we perform receive *antenna selection* [76] over the first hop and transmit antenna selection over the second hop. The relay selection and antenna combining algorithm is presented in the following.

Consider the backward and forward channel matrices of relay k to be written as:

$$\boldsymbol{H}_{\mathcal{SR}_k} = [\boldsymbol{h}_{\mathcal{SR}_k}(1)\ \boldsymbol{h}_{\mathcal{SR}_k}(2)\ \ldots\ \boldsymbol{h}_{\mathcal{SR}_k}(N)]^T \in \mathbb{C}^{N\times M} \tag{4.41}$$

$$\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}} = [\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(1)\ \boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(2)\ \ldots\ \boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(N)]\ \in \mathbb{C}^{M\times N} \tag{4.42}$$

where $\boldsymbol{h}_{\mathcal{SR}_k}(i)^T \in \mathbb{C}^M$ and $\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(j) \in \mathbb{C}^M$ are respectively the channel vectors from source to i^{th} receive antenna of relay $\mathcal{R}_k$ and from relay $\mathcal{R}_k$'s j^{th} transmit antenna to the destination.

1. *Initialization*:

$$\mathcal{T}_0 = \Big\{(\mathcal{R}_k, i, j)\Big|\ |\boldsymbol{h}_{\mathcal{SR}_k}(i)| \geq h_0; |\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(j)| \geq h_0\Big\} \tag{4.43}$$

where h_0 is a threshold value.

2. *Start-up*:

$$(\mathscr{R}_1, b_1, f_1) = \arg \max_{(\mathcal{R}_k,i,j)\in\mathcal{T}_0} \frac{|\boldsymbol{h}_{\mathcal{SR}_k}(j)|^2 |\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(j)|^2}{|\boldsymbol{h}_{\mathcal{SR}_k}(j)|^2 + |\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(j)|^2} \tag{4.44}$$

$$\mathcal{T}_1 = \mathcal{T}_0 - \{(\mathscr{R}_1, b_1, f_1)\} \tag{4.45}$$

$$\Gamma_{(i,j,k)}(1) = \max\Big\{\gamma'_{1,k}(i), \gamma''_{1,k}(j)\Big\} \tag{4.46}$$

where $\gamma'_{1,k}(i)$ and $\gamma''_{1,k}(j)$ are defined as:

$$\gamma'_{1,k}(i) = \varphi\Big(\frac{\boldsymbol{h}_{\mathcal{S}\mathscr{R}_1}(b_1)}{|\boldsymbol{h}_{\mathcal{S}\mathscr{R}_1}(b_1)|}, \frac{\boldsymbol{h}_{\mathcal{SR}_k}(i)}{|\boldsymbol{h}_{\mathcal{SR}_k}(i)|}\Big) \tag{4.47}$$

$$\gamma''_{1,k}(j) = \varphi\Big(\frac{\boldsymbol{h}_{\mathscr{R}_1\mathcal{D}}(f_1)}{|\boldsymbol{h}_{\mathscr{R}_1\mathcal{D}}(f_1)|}, \frac{\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(j)}{|\boldsymbol{h}_{\mathcal{R}_k\mathcal{D}}(j)|}\Big) \tag{4.48}$$

and $\varphi(.,.)$ is defined in (16).

3. *Iterations*: For $2 \leq m \leq M$,

$$(\mathscr{R}_m, b_m, f_m) = \arg \min_{(\mathcal{R}_k,i,j)\in\mathcal{T}_{m-1}} \Gamma_{(i,j,k)}(m-1) \tag{4.49}$$

$$\mathcal{T}_m = \mathcal{T}_{m-1} - \{(\mathscr{R}_m, b_m, f_m)\} \tag{4.50}$$

$$\Gamma_{(i,j,k)}(m) = \Gamma_{(i,j,k)}(m-1) + \max\left\{\gamma'_{m,k}(i), \gamma''_{m,k}(j)\right\}, \tag{4.51}$$

$$\forall(\mathcal{R}_k, i, j) \in \mathcal{T}_m. \tag{4.52}$$

Thus, for the k^{th} mode, the relay $\mathscr{R}_k$ selects the antenna b_k for reception and amplifies the signal by a factor $g_{\mathscr{R}_k}$ and transmits from antenna f_k. Similarly to the eigenmode combining algorithm, the equivalent channel matrix of the first-hop is:

$$\boldsymbol{\mathcal{H}}_{\mathcal{SR}} = \begin{bmatrix} \boldsymbol{h}^T_{\mathcal{S}\mathscr{R}_1}(b_1) \\ \boldsymbol{h}^T_{\mathcal{S}\mathscr{R}_2}(b_2) \\ \vdots \\ \boldsymbol{h}^T_{\mathcal{S}\mathscr{R}_M}(b_M) \end{bmatrix} \in \mathbb{C}^{M\times M} \tag{4.53}$$

and consequently the source beamforming matrix would be:

$$\boldsymbol{W}_{\mathcal{S}} = {\boldsymbol{\mathcal{H}}_{\mathcal{SR}}}^{-1}. \tag{4.54}$$

The equivalent channel matrix of the second hop and the receive beamforming matrix at destination become respectively:

$$\boldsymbol{\mathcal{H}}_{\mathcal{RD}} = \left[\boldsymbol{h}_{\mathscr{R}_1\mathcal{D}}(f_1)|\boldsymbol{h}_{\mathscr{R}_2\mathcal{D}}(f_2)|\ldots|\boldsymbol{h}_{\mathscr{R}_M\mathcal{D}}(f_M)\right] \in \mathbb{C}^{M\times M} \tag{4.55}$$

$$\boldsymbol{W}_{\mathcal{D}} = {\boldsymbol{\mathcal{H}}_{\mathcal{RD}}}^{-1}. \tag{4.56}$$

4.4 Discussion on the Threshold Value

From [77], we know that the sum rate of a MIMO broadcast channel (which can be achieved by dirty paper coding (DPC) [72,78]) scales linearly with the number of transmit antennas and double logarithmically with the number of users [77]. Since we are applying a MIMO multi-user technique at the source node, the best capacity scaling that we may achieve is that of a DPC approach. Therefore, we set the threshold such that the capacity of zero-forcing beamforming (eigenmode-combining or antenna-combining) approaches that of DPC asymptotically or in the other words:

$$\lim_{K\to\infty} C_{DPC} - C_{ZF} = 0. \tag{4.57}$$

Thus, the effective channel gains of each sub-channel from source to relays (and also from relays to the destination) must be comparable to $\log K$. Therefore, we heuristically set the threshold for both eigenmode combining and antenna combining algorithms equal to

$$h_0, \lambda_0 \sim \log K. \tag{4.58}$$

This implies that as the number of relay nodes increases, we set a higher threshold to reduce the amount of feedback from the relays to the source. This threshold is not optimum and the optimum threshold should be a little bit smaller than this threshold.

4.5 Capacity Scaling Law

When it is not possible to exactly characterize the capacity of large networks, a powerful tool is the *capacity scaling* in the *asymptote* of large number of nodes. This powerful tool was introduced in [79][49] for single-antenna AWGN channels and the results have been extended to single-antenna fading channels in [80][81].

Now, we present the capacity scaling of the proposed algorithms. The proof for this theorem is very similar to the proof presented in [70]. The first step is to prove that there exists with probability one a semi-orthogonal subset of relays. The next step is to find the capacity of the selected subset of semi-orthogonal relays and to study the asymptotic behavior of that capacity by considering equal power allocation across each sub-channel.

Theorem 1 *The capacity of eigenmode (antenna) combining algorithm scales as:*

$$C = \frac{M}{2} \log\log K + \mathcal{O}(1) \doteq \frac{M}{2} \log\log K. \tag{4.59}$$

4.6 Simulation Results

In this section, numerical results for the proposed algorithms are presented. In order to give a comprehensive comparison, we use the following systems :

1. *Upper Bound*: By using multi-terminal network information theory and

applying theorem 14.10.1-[24] to two cutset over first hop (broadcast cut) and the second hop (multiple-access cut), we obtain an upper bound on the capacity of the two-hop parallel MIMO relay network. The cut-set around the first hop gives the following rate:

$$R_1 = \max_{\mathbb{P}(\boldsymbol{x}_S)} I(\boldsymbol{x}_S; \boldsymbol{y}_{\mathcal{R}_1}, \boldsymbol{y}_{\mathcal{R}_2}, \ldots, \boldsymbol{y}_{\mathcal{R}_K} | \boldsymbol{x}_{\mathcal{R}_1}, \boldsymbol{x}_{\mathcal{R}_2}, \ldots, \boldsymbol{x}_{\mathcal{R}_K}). \tag{4.60}$$

and the cut-set around the second hop gives us:

$$R_2 = \max_{\mathbb{P}(\boldsymbol{x}_{\mathcal{R}_1}, \boldsymbol{x}_{\mathcal{R}_2}, \ldots, \boldsymbol{x}_{\mathcal{R}_K})} I(\boldsymbol{x}_{\mathcal{R}_1}, \boldsymbol{x}_{\mathcal{R}_1}, \ldots, \boldsymbol{x}_{\mathcal{R}_K}; \boldsymbol{y}_{\mathcal{D}}). \tag{4.61}$$

The upper bound on the capacity is the minimum of these two rate:

$$C_{upper} = \min(R_1, R_2). \tag{4.62}$$

In [13], it is shown that this upper-bound capacity scales as:

$$C_{upper} = \frac{M}{2} \log K + \mathcal{O}(1) \doteq \frac{M}{2} \log K \tag{4.63}$$

2. *Cooperative Beamforming Scheme (CBS)*: This scheme and its variant *Incremental Cooperative Beamforming Scheme (ICBS)* have been proposed in [13]. In this scheme the overall channel matrix from the source node to all relay nodes $\boldsymbol{H}_{\mathcal{SR}}^T = [\boldsymbol{H}_{\mathcal{SR}_1}^T | \boldsymbol{H}_{\mathcal{SR}_2}^T | \ldots | \boldsymbol{H}_{\mathcal{SR}_K}^T]$ is decomposed by singular value decomposition which gives $\boldsymbol{H}_{\mathcal{SR}} = \boldsymbol{U}_{\mathcal{SR}} \boldsymbol{\Lambda}_{\mathcal{SR}}^{\frac{1}{2}} \boldsymbol{V}_{\mathcal{SR}}^*$ where $\boldsymbol{U}_{\mathcal{SR}}$ and $\boldsymbol{V}_{\mathcal{SR}}$ are unitary matrices of sizes $KN \times M$ and $M \times M$ respectively and $\boldsymbol{\Lambda}_{\mathcal{SR}}$ is an $M \times M$ diagonal matrix. $\boldsymbol{U}_{\mathcal{SR}}$ can be par-

titioned into sub-matrices as $[\boldsymbol{U}_{\mathcal{SR}_1}^T|\boldsymbol{U}_{\mathcal{SR}_2}^T|\ldots|\boldsymbol{U}_{\mathcal{SR}_K}^T]$. The k^{th} relay multiplies its received signal by $\boldsymbol{U}_{\mathcal{SR}_k}^*$ and then passes it through a zero-forcing matrix $\boldsymbol{H}_{\mathcal{SR}_k}^\dagger$ and finally amplifies it by a gain factor α which is computed from

$$\alpha = \sqrt{\frac{P_{\mathcal{R}}}{\max_k ||\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}^\dagger \boldsymbol{U}_{\mathcal{SR}_k}^* \boldsymbol{y}_{\mathcal{R}_k}||^2}}. \tag{4.64}$$

The final input-output relation of the channel is given by:

$$\begin{aligned}
\boldsymbol{y}_{\mathcal{D}} &= \alpha \sum_{k=1}^{K} \boldsymbol{H}_{\mathcal{R}_k\mathcal{D}} \boldsymbol{x}_{\mathcal{R}_k} + \boldsymbol{z}_{\mathcal{D}} && (4.65)\\
&= \alpha \sum_{k=1}^{K} \boldsymbol{H}_{\mathcal{R}_k\mathcal{D}} \boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}^\dagger \boldsymbol{U}_{\mathcal{SR}_k}^* \boldsymbol{y}_{\mathcal{R}_k} + \boldsymbol{z}_{\mathcal{D}} && (4.66)\\
&= \alpha \boldsymbol{U}^*(\boldsymbol{H}_{\mathcal{SR}} \boldsymbol{x}_{\mathcal{S}} + \boldsymbol{z}_{\mathcal{R}}) + \boldsymbol{z}_{\mathcal{D}} && (4.67)\\
&= \alpha(\boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{V}^* \boldsymbol{x}_{\mathcal{S}} + \boldsymbol{z}_{\mathcal{R}}^{'}) + \boldsymbol{z}_{\mathcal{D}} && (4.68)
\end{aligned}$$

As we observe the channel is diagonal and the noise vector is white and Gaussian. As the number of relays increases, we expect to have smaller values of α with high probability. In other words, there is a higher chance of having at least one ill-conditioned downlink channel among the relays. In this case, a subset of relays which are in good condition is selected and the rest is turned off. In the ICBS, we select a subset of relays which results in a high value of α. Defining $\beta_k = \mathbb{E}_{\boldsymbol{x}_{\mathcal{S}}, \boldsymbol{z}_{\mathcal{R}_k}}\left[||\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}}^\dagger \boldsymbol{U}_{\mathcal{SR}_k}^* \boldsymbol{y}_{\mathcal{R}_k}||^2\right]$, we activate the relays which satisfy $\beta_k \leq \beta$, where β is a predefined threshold. In this manner, it is guaranteed that $\alpha \geq \sqrt{\frac{P}{\beta}}$. This improvement in the value of α is realized at the

expense of turning off some of the relays, creating interference in the equivalent point-to-point channel. More precisely, by defining $\mathcal{A} = \{k|\beta_k > \beta\}$, we have

$$\boldsymbol{y}_{\mathcal{D}} = \alpha\Big(\Big(\boldsymbol{\Lambda}_{\mathcal{SR}}^{1/2} - \sum_{k\in\mathcal{A}} \boldsymbol{U}_{\mathcal{SR}_k}^{*} \boldsymbol{H}_{\mathcal{SR}_k} \boldsymbol{V}_{\mathcal{SR}}\Big)\boldsymbol{x} + \sum_{k\in\mathcal{A}^c} \boldsymbol{U}_{\mathcal{SR}_k} \boldsymbol{z}_{\mathcal{R}_k}\Big) + \boldsymbol{z}_{\mathcal{D}} \quad (4.69)$$

As the above equation shows, by decreasing the value of β, one can guarantee a large value of α while increasing the gap of the equivalent channel matrix to $\boldsymbol{\Lambda}_{\mathcal{SR}}^{1/2}$. It is shown that by appropriately choosing the value of β , the rate of such a scheme would be at most $\mathcal{O}(\frac{1}{\log K})$ below the corresponding upper bound capacity.

3. *Coherent* relaying: introduced in [10] and [11] where relaying is based on backward and forward zero-forcing at each relay [10] or backward and forward match filtering at every relay node [11]. The capacity of coherent relaying scales as $C_{coherent} = \frac{M}{2} \log K + \mathcal{O}(1) \doteq \frac{M}{2} \log K$.

4. *Non-coherent* relaying or simple amplify-and-forward: where each relay multiplies its received vector by a diagonal matrix $\boldsymbol{G}_k = g\boldsymbol{I}_N$ where g is found from a per-relay long-term power constraint. In this case the capacity scales as $C_{non-coherent} = \frac{M}{2} \log(\mathsf{SNR}) + \mathcal{O}(1) \doteq \frac{M}{2} \log(\mathsf{SNR})$ where SNR is a constant [10], [11].

In Figures (4.3)-(4.12), we have plotted the ergodic capacity of our proposed algorithms together with the four schemes mentioned above. The source and relay transmit powers are assumed to be equal; $P_{\mathcal{S}} = P_{\mathcal{R}} = 0, 10, 20, 30\ dB$ and, $M = 2, 4, 6$. All channels are assumed to be zero-mean unit covariance

Rayleigh fading channel while the effect of large-scale fading is neglected. Mont Carlo simulation are performed over 10000 channel realizations. We can observe the upper-bound capacity linearly scaling in the number of source/destination transmit antennas and logarithmically scaling in the number of relays. The capacity of non-coherent relaying scales like a constant as the number of relays tends to infinity. In fact, increasing the number of relays does not improve the performance of simple amplify-and-forward relaying.

Although coherent relaying array gain is much larger than that of eigenmode combining, due to two-hop interference cancellation by eigenmode combining (compared to distributed interference cancellation in coherent relaying), there is a considerable difference in the capacity of those algorithms for small to medium number of relays. In fact, for small to medium number of relays, eigenmode combining algorithm outperforms coherent relaying [11] and cooperative beamforming [13] because of the superiority of $\mathcal{O}(1)$ term in 4.59 than the $\mathcal{O}(1)$ term in the capacity scaling of coherent relaying and cooperative beamforming. On the other hand, for extremely large number of relays (which is not very common in practice), CBS and coherent relaying achieve higher capacity than the proposed algorithms, due to their asymptotic relaying (logarithmically in K) compared to the proposed algorithms (double-logarithmically in K). In fact we can say that for small to medium number of relays we have outperformed relaying schemes which achieve the asymptotic capacity by sacrificing the array gain and enhancing the $\mathcal{O}(1)$ term. We also observe that the capacity of the antenna combining algorithm is very close to that of eigenmode combining with a slight difference which is due to the extra array gain obtained in eigenmode combining by cooperation

among the antennas. In Figures (4.15)-(4.18), we have plotted the capacity of eigenmode combining algorithm as a function of number of relay nodes for different number antennas at each relay. The simulation results confirm the capacity scaling of eigenmode combining by showing that the value of N does not have any effect on double logarithmic term and only adds up to the $\mathcal{O}(1)$ term.

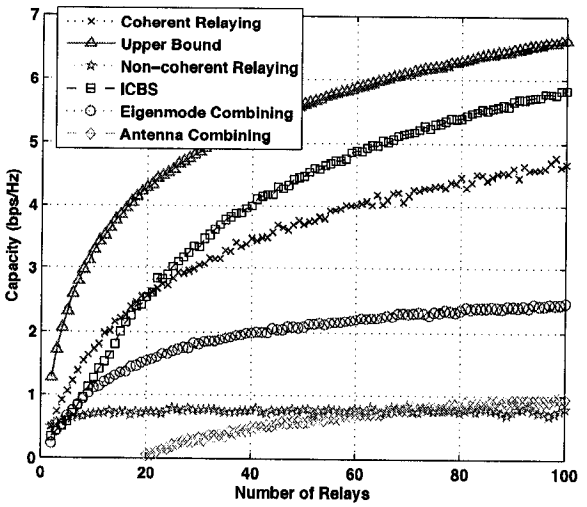

Figure 4.3 Ergodic capacity as a function of number of relays for $M = N = 2$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 0\ dB$.

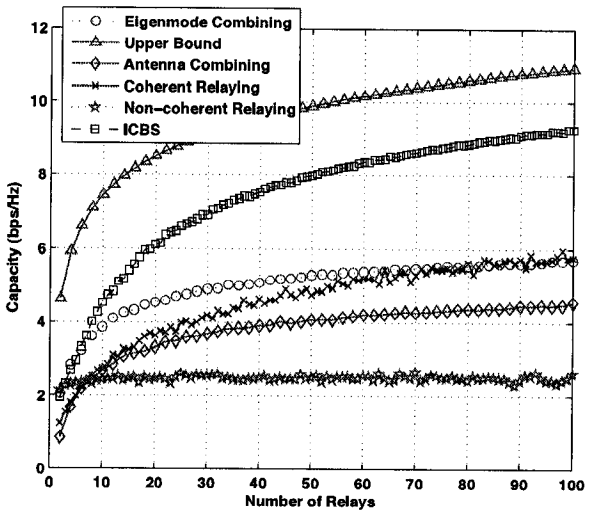

Figure 4.4 Ergodic capacity as a function of number of relays for $M = N = 2$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 10\ dB$.

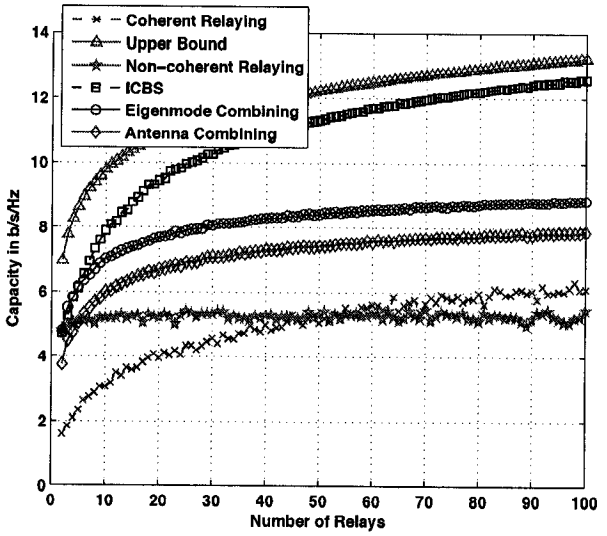

Figure 4.5 Ergodic capacity as a function of number of relays for $M = N = 2$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 20\ dB$.

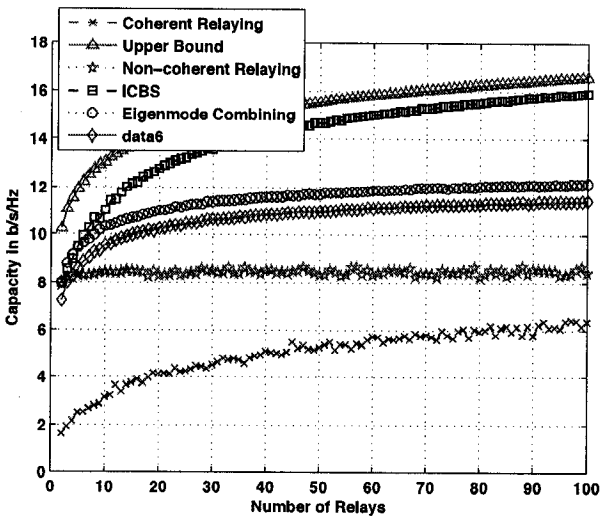

Figure 4.6 Ergodic capacity as a function of number of relays for $M = N = 2$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 30\ dB$.

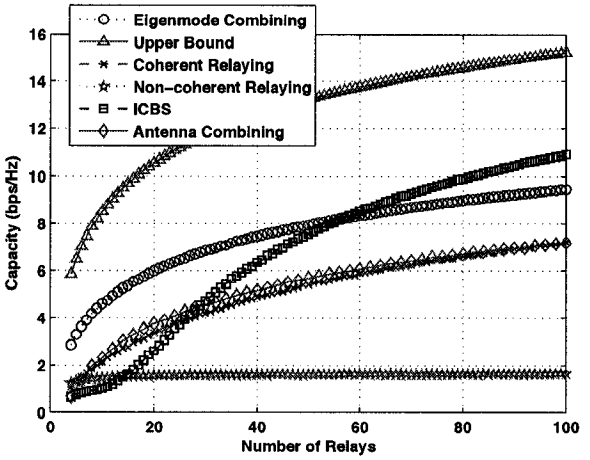

Figure 4.7 Ergodic capacity as a function of number of relays for $M = N = 4$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 0\ dB$.

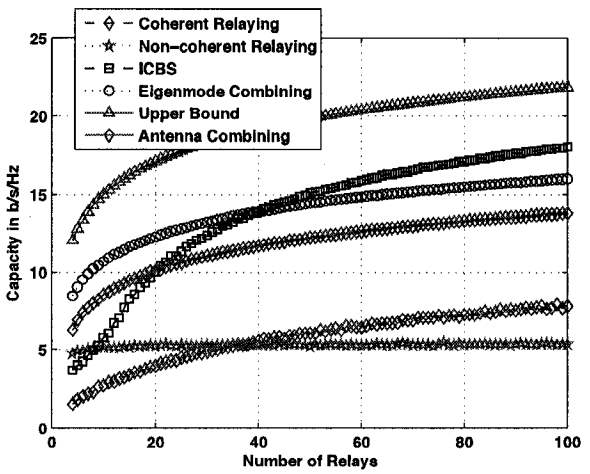

Figure 4.8 Ergodic capacity as a function of number of relays for $M = N = 4$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 10\ dB$.

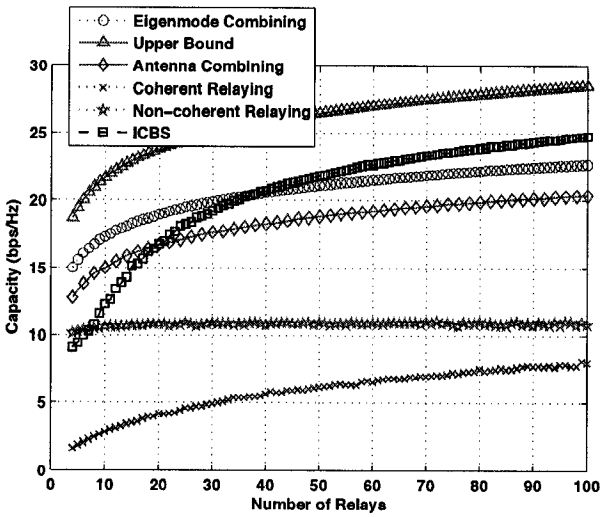

Figure 4.9 Ergodic capacity as a function of number of relays for $M = N = 4$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 20\ dB$.

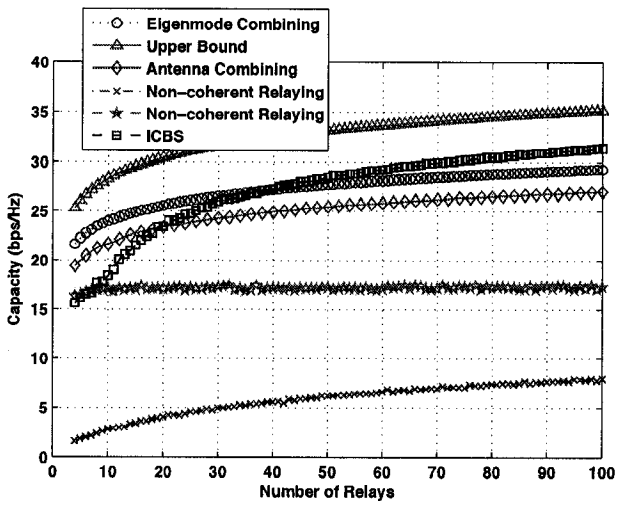

Figure 4.10 Ergodic capacity as a function of number of relays for $M = N = 4$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 30\ dB$.

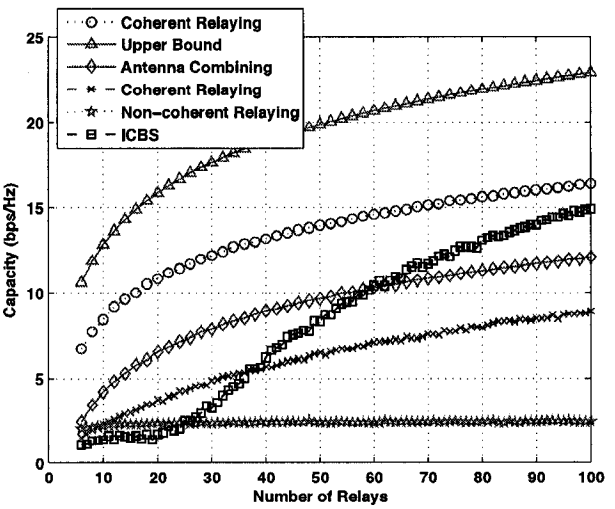

Figure 4.11 Ergodic capacity as a function of number of relays for $M = N = 6$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 0\ dB$.

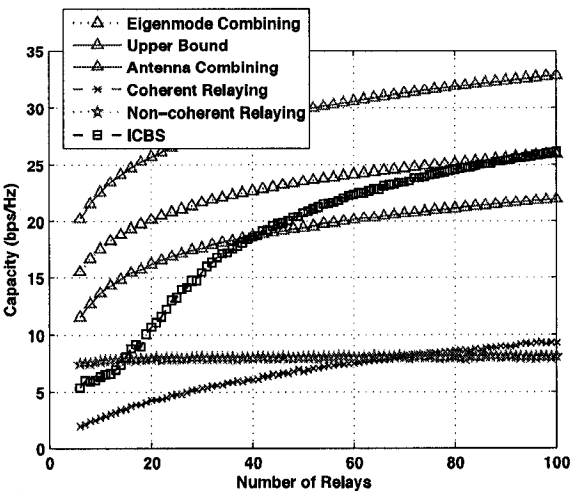

Figure 4.12 Ergodic capacity as a function of number of relays for $M = N = 6$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 10\ dB$.

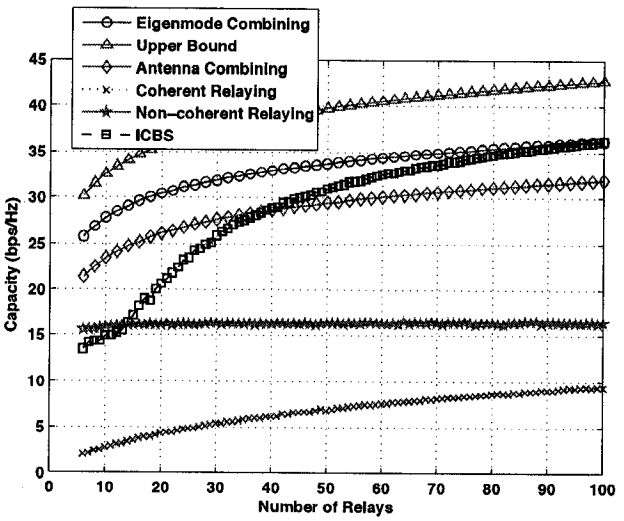

Figure 4.13 Ergodic capacity as a function of number of relays for $M = N = 6$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 20\ dB$.

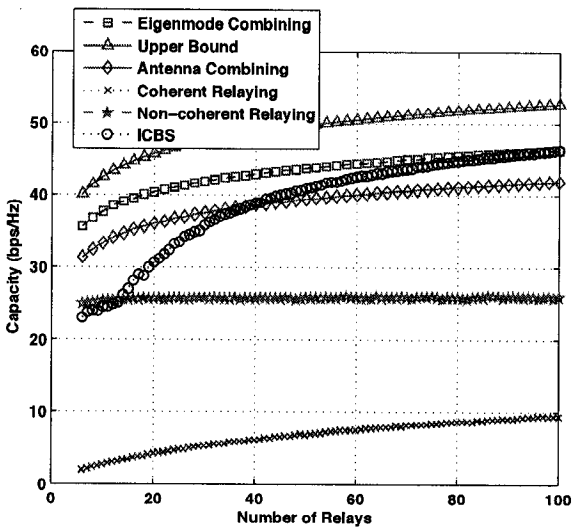

Figure 4.14 Ergodic capacity as a function of number of relays for $M = N = 6$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 30\ dB$.

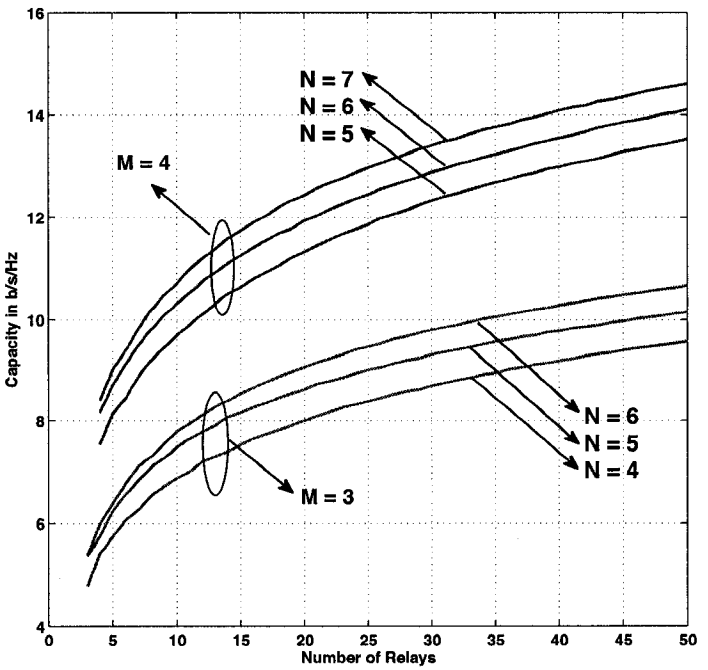

Figure 4.15 Ergodic capacity (Eigenmode Combining) as a function of number of relays K and the number of antennas at each relay N for $M = 3, 4$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 10\ dB$.

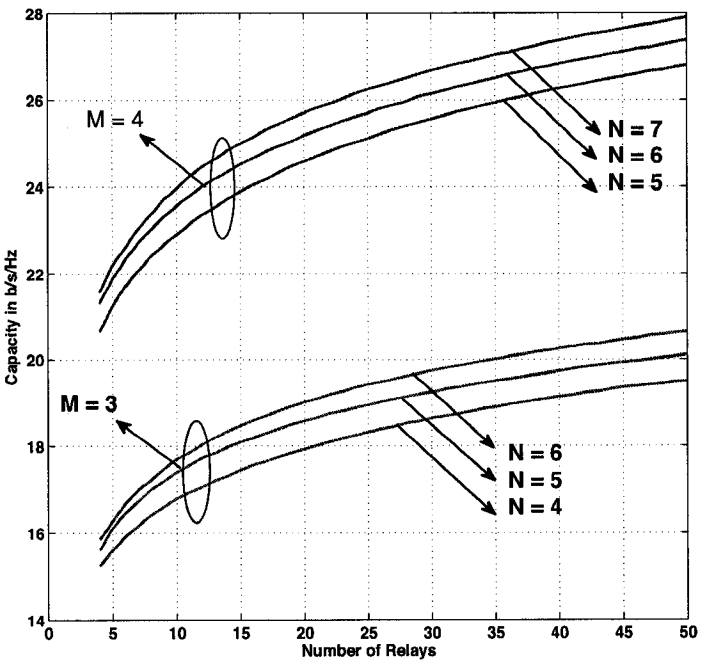

Figure 4.16 Ergodic capacity (Eigenmode Combining) as a function of number of relays K and the number of antennas at each relay N for $M = 3, 4$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 30\ dB$.

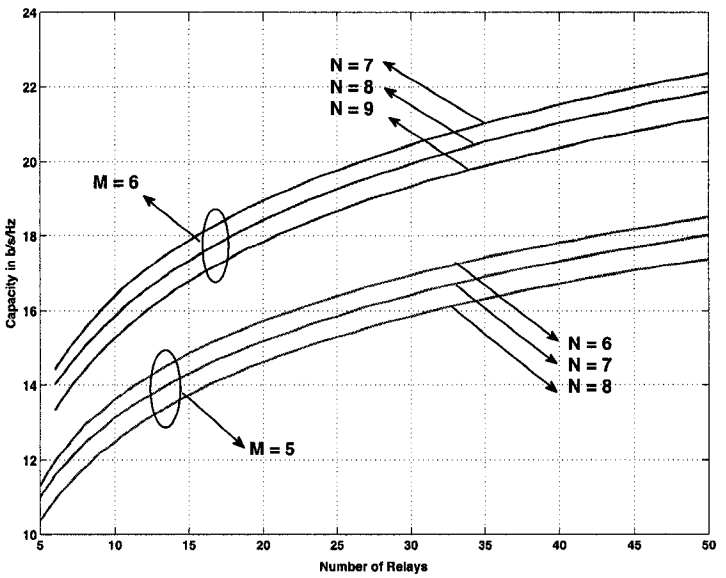

Figure 4.17 Ergodic capacity (Eigenmode Combining) as a function of number of relays K and the number of antennas at each relay N for $M = 5, 6$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 10\ dB$.

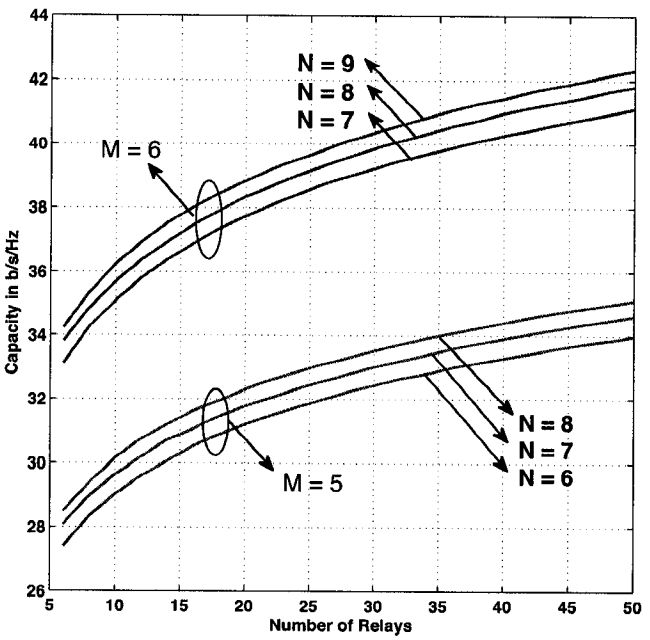

Figure 4.18 Ergodic capacity (Eigenmode Combining) as a function of number of relays K and the number of antennas at each relay N for $M = 5, 6$ and $P_{\mathcal{S}} = P_{\mathcal{R}} = 30\ dB$.

CHAPTER 5

MIMO INTERFERENCE RELAY NETWORKS

5.1 Introduction

Noise, fading and interference are the three communications bottlenecks that have challenged engineers since the early days of wireless communications. Although we can argue that our knowledge of noise and fading mitigation techniques is mature enough, the same argument is not valid for interference. It was not until mid 80's that the optimistic view on interference started to flourish especially for centralized networks [82]. Before that, it was assumed that interference can not be mitigated and at best we can treat interference as AWGN. However, despite these advances on mitigating interference in centralized networks, little is known about the distributed interference channels (IC) where multiple source nodes communicate with their paired destination nodes over a shared physical channel.

In the previous chapter, we considered a scenario where a single multi-antenna source-destination pair communicate via several multi-antenna relay terminals located in the region between the source and the destination. An extension to this scenario is where there are multiple source-destination pairs communicating via several relays over the same common physical channel. This configuration is known in the literature as *MIMO interference relay net-*

work (MIRN) [12]. The complexity of this configuration stems from the fact that there are not only *inter-stream interferences (ISI)* over each source-destination link, but also there are *inter-user interferences (IUI)* between different source-destination pairs. Besides that, because of the distributed nature of the network, cooperation at source, relay and destination level is very difficult.

In order to address these problems and motivate our approach, let us first define the interference channel and then review the background works. Interference channel (IC) is a long-standing problem which was first treated by Shannon [83]. It models the situations where multiple *non-cooperative* source-destination pairs communicate over the same physical channel. The two-user discrete memoryless interference channel is defined by the conditional probability $\mathbb{P}(y_1, y_2|x_1, x_2)$ where $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$, $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$, $\mathcal{X}_1$ and $\mathcal{X}_2$ are the finite sets of inputs, and $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are the finite sets of outputs. For coded information of block length N, the two-user interference channel is denoted by $(\mathcal{X}_1^N \times \mathcal{X}_2^N, \mathbb{P}^N(\boldsymbol{y}_1, \boldsymbol{y}_2|\boldsymbol{x}_1, \boldsymbol{x}_2), \mathcal{Y}_1^N \times \mathcal{Y}_2^N)$ where

$$\mathbb{P}^N(\boldsymbol{y}_1, \boldsymbol{y}_2|\boldsymbol{x}_1, \boldsymbol{x}_2) = \prod_{n=1}^{N} \mathbb{P}(y_1(n), y_2(n)|x_1(n), x_2(n)),$$
$$\boldsymbol{x}_1 \in \mathcal{X}_1^N, \boldsymbol{x}_2 \in \mathcal{X}_2^N, \boldsymbol{y}_1 \in \mathcal{Y}_1^N, \boldsymbol{y}_1 \in \mathcal{Y}_1^N. \tag{5.1}$$

The two independent non-cooperative sources produce two integers $W_1 \in \{1, \ldots, 2^{NR_1}\}$, $W_2 \in \{1, \ldots, 2^{NR_2}\}$, respectively. A $(2^{NR_1}, 2^{NR_2}, N)$ code for

a two-user interference channel consists of two encoding functions

$$e_1 \ : \ \{1,\ldots,2^{NR_1}\} \rightarrow \mathcal{X}_1^N \tag{5.2}$$

$$e_2 \ : \ \{1,\ldots,2^{NR_2}\} \rightarrow \mathcal{X}_2^N \tag{5.3}$$

and two decoding functions

$$d_1 \ : \ \mathcal{Y}_1^N \rightarrow \{1,\ldots,2^{NR_1}\} \tag{5.4}$$

$$d_2 \ : \ \mathcal{Y}_2^N \rightarrow \{1,\ldots,2^{NR_2}\} \tag{5.5}$$

We say a rate pair (R_1, R_2) is achievable if there exists at least a sequence of $(2^{\lceil NR_1 \rceil}, 2^{\lceil NR_2 \rceil}, N)$ codes such that $\mathbb{P}_{e.1} \rightarrow 0$ and $\mathbb{P}_{e,2} \rightarrow 0$ as $N \rightarrow \infty$ where $\mathbb{P}_{e,1}$ and $\mathbb{P}_{e,2}$ denote the average probability of error for source-destination pairs 1 and 2. The *capacity region* of the interference channel is defined as the closure of the set of all achievable rate pairs (R_1, R_2).

The capacity region of Gaussian interference channels is in general unknown yet, except for very strong interference cases [84,84–87]. The first formulation of capacity region was reported by Shannon in [83] which was computationally prohibitive. In a seminal work by Carleial [88], sequential superposition coding [89] was proposed to achieve a new lower bound for interference channels. In a later work [85], Han and Kobayashi applied simultaneous superposition coding to obtain a larger lower bound which remains to be the best reported to date. There are also few upper bounds on the capacity of interference channel in the literature [90–92]. The capacity region of K-user interference channel has also been recently under investigation. Two recent

works can be found in [93, 94].

The conventional conjecture on the distributed interference networks was that because of interfering sources, at most one degree of freedom can be achieved [57]. Instead of avoiding interfering sources, the effect of interfering signals may be mitigated by active scatterers using amplify-and-forward (AF) relaying. Bolcskei, *et al* took this approach namely, "*distributed* interference cancellation" by placing multiple relay nodes between the source and destination nodes to orthogonalize the IC [10, 11]. By simple backward and forward zero-forcing or matched-filtering at the relay nodes, they improved the degrees of freedom of interference networks in the *asymptote* of large number of relay nodes. Parallel to that work, it was proved in [95] that multi-hop AF relaying does not reduce the degrees of freedom as long as there is full cooperation at the source and destination level. For finite number of single-antenna relays, a scheme which is based on zero-forcing by performing a nullspace projection is presented in [63]. The gain factors of the AF relays are chosen such that interference at the destination nodes is cancelled. A global phase reference (partial relay cooperation) as well as channel knowledge are required at the relays.

Extending the results of interference channel to MIMO systems has been less well studied. Viswanath and Jafar investigated the case where either transmitters or receivers are equipped with multiple antennas [96]. In [97], the superposition coding technique is utilized to derive an inner-bound for the capacity of the MIMO interference channels. Blum considered the capacity region and optimal signaling for MIMO interference channel with CSIT in

[98] and with no CSIT in [99]. In [100] the multiplexing gain of the MIMO interference channel with arbitrary number of transmit and receive antennas is derived by Jafar and Fakhreddin. All these works have assumed that there is *no* cooperation between either transmitters or receivers. In [101, 102] it is shown that if there exists an infinite-rate error-free link between transmitters or receivers (i.e. *perfect cooperation*), we can further increase the multiplexing gain of MIMO interference channels. If the infinite-rate error-free link is replaced by a fading noisy channel, the multiplexing gain of the interference channel reduces to one [103]. The result in [103] together with [57] are in a sense negative since they show that cooperation over fading noisy channels does *not* increase the multiplexing gain of the system. Even in the case of full cooperation between transmitters, there is no proven out-performance over TDMA or FDMA approaches in terms of multiplexing gain.

In [104–106], a more general case of IC known as X channel was envisioned where every transmitting node has a message for every receiving node. This channel integrates IC, broadcast (BC) and multiple-access (MAC) channels. Maddah-Ali *et al* showed that by using some spatial filters, this channel can be decomposed into two non-interfering BC or MAC. The optimality of their signaling in terms of achieving integer degrees of freedom has been proven in [107]. The degrees of freedom region of an X channel with two source and two destination nodes all equipped with multiple antenna was found in [108] by Jafar and Shamai. There, they used the *relativity* of interference and the concept of *interference alignment* to achieve the limits of MIMO X channel in terms of degrees of freedom. Relativity of interference means that the same signal that is favorable to a user may be viewed as undesirable for other

users and hence, interference depends totally on the signal-space view point of receivers. Therefore, interference alignment techniques try to force interfering source to cast overlapping interferences at the undesired destinations to save in the dimension of interference space and in return, increase the dimension of signal space at desired receivers. Interference alignment techniques have also found applications in the multi-user IC [109] and wireless X networks [93]. Despite these marvelous features, interference alignment requires that all the transmitting and receiving nodes have *global* channel state information (CSI) of the network which is difficult to obtain especially in large networks with multi-antenna nodes. It also needs multiple (long) extensions of the channel over time or frequency which is not feasible in some scenarios.

In this chapter, we consider a two-hop MIMO interference relay network (MIRN) as in [11]. Similar to [110], we regard the first and second hop as two cascaded MIMO X channels since X channels provide higher degrees of freedom than IC. Interference alignment and the integer degrees-of-freedom-region is discussed when there is no channel extension over time or frequency and the nodes need not to have global CSI of the network. We will jointly optimize the degrees-of freedom region and beamfoming vectors to adapt to channel conditions. By doing so, we will decompose this channel into multiple non-interfering parallel MIMO relay channels as described in [14]. We will show that by sacrificing the total degrees of freedom to enhance "interference alignment", we are able to achieve higher capacity for small to medium number of relay terminals.

The rest of this chapter is organized as follows: in section 5.2, we present

the model for a simple MIMO interference relay network with two sources, two destinations and two relay nodes. In section 5.3, we discuss the decomposition approach to MIMO interference relay networks. In section 5.4, we extend the system model described in section 5.3 to a larger network with arbitrary number of source-destination pairs and relay nodes. In section 5.5, we formulate the degrees-of-freedom maximization problem and present the decomposition approach. In section 5.6, sub-optimum signaling and interference alignment for MIMO interference relay network is discussed. Simulation results and conclusions will follow in sections 5.7 and 5.8, respectively.

5.2 Network Model

For simplicity of the exposition, we consider a MIMO interference relay network with two source-destination pairs denoted by $(\mathcal{S}_1,\mathcal{D}_1)$ and $(\mathcal{S}_2,\mathcal{D}_2)$ communicating via two relay nodes $\mathcal{R}_1$ and $\mathcal{R}_2$ which are located in the middle region between the sources and the destinations (see Figure 5.1). Extension of this model to a larger network is presented in section 5.4. We make the following assumptions on the number of antennas at each node $i = 1, 2$:

$$M_{\mathcal{S}_i} = M_{\mathcal{D}_i} = M_i, \tag{5.6}$$

$$M_{\mathcal{R}_i} = N_i. \tag{5.7}$$

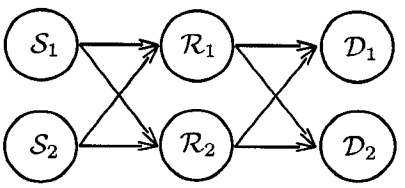

Figure 5.1 Block diagram of a two-hop MIMO interference relay network with two source-destination pairs and two relay nodes.

There are also some constraints on the values of M_1, M_2, N_1 and N_2 as follows:

$$M_1 \geq M_2, N_1 \geq N_2, \tag{5.8}$$

$$M_1 < N_1 + N_2, \tag{5.9}$$

$$N_1 < M_1 + M_2. \tag{5.10}$$

Condition (5.8) is made for simplifying the problem. Conditions (5.9) and (5.10) are justified by the fact that if $M_1 \geq N_1+N_2$ or $N_1 \geq M_1+M_2$, then the multiplexing gain is restricted to N_1+N_2 and M_1+M_2, respectively per each transmission in time and frequency. Data transmission from $\mathcal{S}_1$ and $\mathcal{S}_2$ to $\mathcal{D}_1$ and $\mathcal{D}_2$ is performed according to the well-known amplify-and-forward (AF) half-duplex relaying protocol discussed in the previous chapters. Assuming perfect synchronization between $\mathcal{S}_1$ and $\mathcal{S}_2$, we have the following received

signals at the relay nodes over the first time slot:

$$\boldsymbol{y}_{\mathcal{R}_1} = \boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_1}\boldsymbol{x}_{\mathcal{S}_1} + \boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_1}\boldsymbol{x}_{\mathcal{S}_2} + \boldsymbol{z}_{\mathcal{R}_1}, \tag{5.11}$$

$$\boldsymbol{y}_{\mathcal{R}_2} = \boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2}\boldsymbol{x}_{\mathcal{S}_1} + \boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2}\boldsymbol{x}_{\mathcal{S}_2} + \boldsymbol{z}_{\mathcal{R}_2} \tag{5.12}$$

where $\boldsymbol{x}_{\mathcal{S}_i}$ denotes the transmit signal from source node $i, i = 1, 2$; under the following power constraint:

$$Tr(\mathbb{E}\{\boldsymbol{x}_{\mathcal{S}_1}\boldsymbol{x}_{\mathcal{S}_1}^*\}) \leq P_{\mathcal{S}_1} \tag{5.13}$$

$$Tr(\mathbb{E}\{\boldsymbol{x}_{\mathcal{S}_2}\boldsymbol{x}_{\mathcal{S}_2}^*\}) \leq P_{\mathcal{S}_2} \tag{5.14}$$

$$P_{\mathcal{S}_1} + P_{\mathcal{S}_2} = P_{\mathcal{S}}, \tag{5.15}$$

$\boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_j}$ denotes the $(M_{\mathcal{D}_i} \times M_{\mathcal{R}_j})$ channel matrix corresponding to $\mathcal{S}_i \rightarrow \mathcal{R}_j$ link and finally $\boldsymbol{z}_{\mathcal{R}_j} \in \mathbb{C}^{M_{\mathcal{R}_j} \times 1}$ is the unit variance ZMCSCG noise vector at relay $\mathcal{R}_j; j = 1, 2$. In the second time slot, relay $\mathcal{R}_j$ process its received vector $\boldsymbol{y}_{\mathcal{R}_j}$ by a processing matrix $\boldsymbol{G}_{\mathcal{R}_j} \in \mathbb{C}^{M_{\mathcal{R}_j} \times M_{\mathcal{R}_j}}$ under the following power constraints:

$$Tr(\mathbb{E}\{\boldsymbol{x}_{\mathcal{R}_1}\boldsymbol{x}_{\mathcal{R}_1}^*\}) = Tr(\mathbb{E}\{(\boldsymbol{G}_{\mathcal{R}_1}\boldsymbol{y}_{\mathcal{R}_1})(\boldsymbol{G}_{\mathcal{R}_1}\boldsymbol{y}_{\mathcal{R}_1})^*\}) \leq P_{\mathcal{R}_1}, \tag{5.16}$$

$$Tr(\mathbb{E}\{\boldsymbol{x}_{\mathcal{R}_2}\boldsymbol{x}_{\mathcal{R}_2}^*\}) = Tr(\mathbb{E}\{(\boldsymbol{G}_{\mathcal{R}_1}\boldsymbol{y}_{\mathcal{R}_1})(\boldsymbol{G}_{\mathcal{R}_1}\boldsymbol{y}_{\mathcal{R}_1})^*\}) \leq P_{\mathcal{R}_2}, \tag{5.17}$$

$$P_{\mathcal{R}_1} + P_{\mathcal{R}_2} \leq P_{\mathcal{R}}. \tag{5.18}$$

Assuming perfect synchronization between $\mathcal{R}_1$ and $\mathcal{R}_2$, we obtain the follow-

ing signals at the destination nodes $\mathcal{D}_1$ and $\mathcal{D}_2$ over the second time slot:

$$\boldsymbol{y}_{\mathcal{D}_1} = \boldsymbol{H}_{\mathcal{R}_1\mathcal{D}_1}\boldsymbol{x}_{\mathcal{R}_1} + \boldsymbol{H}_{\mathcal{R}_2\mathcal{D}_1}\boldsymbol{x}_{\mathcal{R}_2} + \boldsymbol{z}_{\mathcal{D}_1}, \tag{5.19}$$

$$\boldsymbol{y}_{\mathcal{D}_2} = \boldsymbol{H}_{\mathcal{R}_1\mathcal{D}_2}\boldsymbol{x}_{\mathcal{R}_1} + \boldsymbol{H}_{\mathcal{R}_2\mathcal{D}_2}\boldsymbol{x}_{\mathcal{R}_2} + \boldsymbol{z}_{\mathcal{D}_2}, \tag{5.20}$$

where $\boldsymbol{H}_{\mathcal{R}_j\mathcal{D}_i}$ denotes the $(M_{\mathcal{R}_j} \times M_{\mathcal{S}_i})$ channel matrix corresponding to $\mathcal{R}_j \rightarrow \mathcal{D}_i$ link $\forall i, j \in \{1, 2\}$, and finally $\boldsymbol{z}_{\mathcal{D}_j} \in \mathbb{C}^{M_{\mathcal{D}_i} \times 1}$ is the unit variance ZMCSCG noise vector at destination node $\mathcal{D}_i$, $i = 1, 2$.

Definition 1 *A (M_1, M_2, N_1, N_2) MIMO interference relay network that satisfies conditions (5.8),(5.9),(5.10) is called* irreducible, *if*

$$N_1 \geq N_2 \geq M_1 \geq M_2. \tag{5.21}$$

Otherwise, it is called reducible *to an irreducible MIMO interference relay network.*

Unlike the irreducible systems, a portion of the achieved multiplexing gain in a reducible channel is attained through exploiting the null-spaces of the direct or cross links. In the reducible systems, the null-spaces of the links provide the possibility to increase the number of data streams sent from one of the transmitters to one of the receivers, without imposing any interference on the other receiver or restricting the signaling space of the other transmitter. By excluding the null spaces to increase the multiplexing gain, the system is reduced to an equivalent system with (M_1', M_2', N_1', N_2') antennas, where $(M_1', M_2', N_1', N_2') \leq (M_1, M_2, N_1, N_2)$. As will be explained later, the null-

spaces of the links in the reducible systems are exploited to the extend that the equivalent (reduced) system is not reducible anymore.

We define $\mu'_{\mathcal{S}_i\mathcal{R}_j}$ and $\mu'_{\mathcal{R}_i\mathcal{D}_j}$, $i,j = 1,2$, as the number of data streams sent from $\mathcal{S}_i$ to $\mathcal{R}_j$ and $\mathcal{R}_i$ to $\mathcal{D}_j$, respectively, excluding the number of data streams obtained through null spaces of the links. On the other hand, $\mu_{\mathcal{S}_i\mathcal{R}_j}$ and $\mu_{\mathcal{R}_i\mathcal{D}_j}$ represent the total number of data streams in MIRN.

5.3 Decomposition Approach

In order to obtain the maximum total degrees of freedom, we regard the first and second hop as two cascaded MIMO X channels. The rationale behind this is that MIMO X channels provide degrees of freedom close to that of point-to-point MIMO channel. For example, a point-to point MIMO channel with M transmit and N receive antennas provide $\min\{M,N\}$ degrees of freedom [4] while an X channel with M distributed transmit and N distributed receive antennas provide $\frac{MN}{M+N-1}$ degrees of freedom [93]. The difference between these two figures vanishes as either M or N converge to infinity. In the case of MIRN, as the number of relay terminals increases, we expect that the degrees of freedom of the network will approach that of a point-to-point MIMO channel. In this section, we apply the signaling technique for MIMO X channels as described in [104–106, 111] into the first and second hop of MIRN to justify the techniques that we introduce in section 5.6. First, we give an overview of the definition of degrees of freedom region for MIRN described in the previous section and then we will present the degrees of

freedom maximization problem as an integer linear programming problem.

To do so, multiple transmit and receive filters are utilized at every transmitting and receiving node . More specifically, at source node $\mathcal{S}_i$ the transmit filter $\boldsymbol{T}_{\mathcal{S}_i} \in \mathbb{OC}^{M_i \times (\mu_{i1}+\mu_{i2})}$ is used to multiplex μ_{i1} data streams to $\mathcal{R}_1$ and μ_{i2} data streams to $\mathcal{R}_2$, where $\mathbb{OC}^{M\times N}$ denotes the set of all $M \times N$ complex matrices with mutually orthogonal and normal columns. The transmit filters $\boldsymbol{T}_{\mathcal{S}_i}$, $i = 1, 2$, have two functionalities:

1. confining the transmit signal from $\mathcal{S}_i$ to a $(\mu_{i1} + \mu_{i2})$-dimensional subspace which provides the possibility of decomposing the system into two MIMO parallel relay channels by using linear filters,

2. exploiting the null spaces of the channel matrices to achieve the highest multiplexing gain.

At each relay, two parallel receive and two parallel transmit filters are employed. The received vector $\boldsymbol{y}_{\mathcal{R}_1}$ at $\mathcal{R}_1$ is passed through the filter $\boldsymbol{R}^*_{\mathcal{S}_1\mathcal{R}_1}$, which is used to null out the signal coming from $\mathcal{S}_2$. The μ_{11} data streams sent by $\mathcal{S}_1$ intended to $\mathcal{R}_1$ can be extracted from the output of the receive filter $\boldsymbol{R}^*_{\mathcal{S}_1\mathcal{R}_1}$. After extracting these data streams, $\mathcal{R}_1$ amplifies the extracted data streams by the linear matrix $\boldsymbol{G}_{\mathcal{S}_1\mathcal{R}_1} \in \mathbb{C}^{\mu_{11}\times\mu_{11}}$ and forwards the amplified data streams to $\mathcal{D}_1$ by passing through the filter $\boldsymbol{T}^*_{\mathcal{R}_1\mathcal{D}_1}$. Similarly, to extract μ_{21} data streams, sent by $\mathcal{S}_2$ to $\mathcal{R}_1$, the received vector $\boldsymbol{y}_{\mathcal{R}_1}$ is passed through the receive filter $\boldsymbol{R}^*_{\mathcal{S}_2\mathcal{R}_1}$ (which is used to null out the signal coming from $\mathcal{S}_1$), amplified by the linear gain matrix $\boldsymbol{G}_{\mathcal{S}_2\mathcal{R}_1} \in \mathbb{C}^{\mu_{21}\times\mu_{21}}$ and passed through transmit filter $\boldsymbol{T}^*_{\mathcal{R}_1\mathcal{D}_2}$. $\mathcal{R}_2$ has the same structure as $\mathcal{R}_1$.

Finally, the destination nodes $\mathcal{D}_1$ and $\mathcal{D}_2$ employ receive filters $\boldsymbol{R}_{\mathcal{D}_1}$ and $\boldsymbol{R}_{\mathcal{D}_2}$ to respectively demultiplex the data streams received from the relays $\mathcal{R}_1$ and $\mathcal{R}_2$. Later, it is shown that if μ_{ij}, $\forall i, j \in \{1, 2\}$, satisfy a set of inequalities, then it is possible to design the transmit and receive filters to meet the desired features explained earlier. It means that the system is decomposed into two *non-interfering* parallel MIMO relay channels (see Figure 5.2).

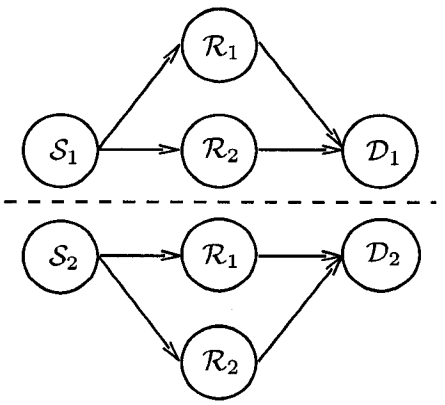

Figure 5.2 Decomposition of a two-hop MIMO interference relay network into two non-interfering parallel MIMO relay channels.

Next, we explain how to select the design parameters including the number of data streams μ_{ij}, $\forall i, j \in \{1, 2\}$ and the transmit-receive filters. The primary objective is to *maximize* the total degrees of freedom of the network, i.e. $\mu_{11} + \mu_{12} + \mu_{21} + \mu_{22}$. The integer variables $\zeta_{\mathcal{S}_i\mathcal{R}_j}, \zeta_{\mathcal{R}_i\mathcal{D}_j}$; $\forall i, j \in \{1, 2\}$, defined below, will be useful in the subsequent discussions (Note that $\Omega(\boldsymbol{A})$ denotes the sub-space spanned by columns of $\boldsymbol{A}$.):

- $\zeta_{\mathcal{S}_1\mathcal{R}_1}$ and $\zeta_{\mathcal{R}_1\mathcal{D}_1}$ denote the dimensions of $\Omega(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_1}\boldsymbol{T}_{\mathcal{S}_2})$ and $\Omega(\boldsymbol{H}^*_{\mathcal{R}_1\mathcal{D}_2}\boldsymbol{R}_{\mathcal{D}_2})$

respectively,

- $\zeta_{\mathcal{S}_1\mathcal{R}_2}$ and $\zeta_{\mathcal{R}_2\mathcal{D}_1}$ denote the dimensions of $\Omega(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2}\boldsymbol{T}_{\mathcal{S}_2})$ and $\Omega(\boldsymbol{H}^*_{\mathcal{R}_1\mathcal{D}_1}\boldsymbol{R}_{\mathcal{S}_1})$ respectively,
- $\zeta_{\mathcal{S}_2\mathcal{R}_1}$ and $\zeta_{\mathcal{R}_1\mathcal{D}_2}$ denote the dimensions of $\Omega(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_1}\boldsymbol{T}_{\mathcal{S}_1})$ and $\Omega(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_2}\boldsymbol{R}_{\mathcal{D}_2})$ respectively,
- $\zeta_{\mathcal{S}_2\mathcal{R}_2}$ and $\zeta_{\mathcal{R}_2\mathcal{D}_2}$ denote the dimensions of $\Omega(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2}\boldsymbol{T}_{\mathcal{S}_1})$ and $\Omega(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_1}\boldsymbol{R}_{\mathcal{S}_1})$ respectively.

In the sequel, we only treat the case of an irreducible system (Case I). The other cases (Case II-IV) are explained in appendix I. To facilitate the derivations, we use the auxiliary variables $(M_1^{'}, M_2^{'}, N_1^{'}, N_2^{'})$. As will be explained later, for each case, $(M_1^{'}, M_2^{'}, N_1^{'}, N_2^{'})$ are computed directly as a function of (M_1, M_2, N_1, N_2). $(\mu^{'}_{\mathcal{S}_1\mathcal{R}_1}, \mu^{'}_{\mathcal{S}_1\mathcal{R}_2}, \mu^{'}_{\mathcal{S}_2\mathcal{R}_1}, \mu^{'}_{\mathcal{S}_2\mathcal{R}_2})$ and $(\mu^{'}_{\mathcal{R}_1\mathcal{D}_1}, \mu^{'}_{\mathcal{R}_1\mathcal{D}_2}, \mu^{'}_{\mathcal{R}_2\mathcal{D}_1}, \mu^{'}_{\mathcal{D}_2\mathcal{D}_2})$ are then selected such that the following constraints are satisfied:

$$
\begin{aligned}
&\mu^{'}_{\mathcal{S}_1\mathcal{R}_1} : \mu^{'}_{\mathcal{S}_1\mathcal{R}_1} + \mu^{'}_{\mathcal{S}_2\mathcal{R}_1} + \mu^{'}_{\mathcal{S}_2\mathcal{R}_2} \leq N_1^{'} \quad , \quad \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} : \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} + \mu^{'}_{\mathcal{R}_2\mathcal{D}_1} + \mu^{'}_{\mathcal{R}_2\mathcal{D}_2} \leq N_1^{'},\\
&\mu^{'}_{\mathcal{S}_2\mathcal{R}_1} : \mu^{'}_{\mathcal{S}_2\mathcal{R}_1} + \mu^{'}_{\mathcal{S}_1\mathcal{R}_1} + \mu^{'}_{\mathcal{S}_1\mathcal{R}_2} \leq N_1^{'} \quad , \quad \mu^{'}_{\mathcal{R}_2\mathcal{D}_1} : \mu^{'}_{\mathcal{R}_2\mathcal{D}_1} + \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} + \mu^{'}_{\mathcal{R}_1\mathcal{D}_2} \leq N_1^{'},\\
&\mu^{'}_{\mathcal{S}_2\mathcal{R}_2} : \mu^{'}_{\mathcal{S}_2\mathcal{R}_2} + \mu^{'}_{\mathcal{S}_1\mathcal{R}_2} + \mu^{'}_{\mathcal{S}_1\mathcal{R}_1} \leq N_2^{'} \quad , \quad \mu^{'}_{\mathcal{R}_2\mathcal{D}_2} : \mu^{'}_{\mathcal{R}_2\mathcal{D}_2} + \mu^{'}_{\mathcal{R}_1\mathcal{D}_2} + \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} \leq N_2^{'},\\
&\mu^{'}_{\mathcal{S}_1\mathcal{R}_2} : \mu^{'}_{\mathcal{S}_1\mathcal{R}_2} + \mu^{'}_{\mathcal{S}_2\mathcal{R}_2} + \mu^{'}_{\mathcal{S}_2\mathcal{R}_1} \leq N_2^{'} \quad , \quad \mu^{'}_{\mathcal{R}_1\mathcal{D}_2} : \mu^{'}_{\mathcal{R}_1\mathcal{D}_2} + \mu^{'}_{\mathcal{R}_2\mathcal{D}_2} + \mu^{'}_{\mathcal{R}_2\mathcal{D}_1} \leq N_2^{'},\\
&\qquad\qquad \mu^{'}_{\mathcal{S}_1\mathcal{R}_1} + \mu^{'}_{\mathcal{S}_1\mathcal{R}_2} \leq M_1^{'} \quad , \quad \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} + \mu^{'}_{\mathcal{R}_1\mathcal{D}_2} \leq N_1^{'},\\
&\qquad\qquad \mu^{'}_{\mathcal{S}_2\mathcal{R}_2} + \mu^{'}_{\mathcal{S}_2\mathcal{R}_1} \leq M_2^{'} \quad , \quad \mu^{'}_{\mathcal{R}_2\mathcal{D}_2} + \mu^{'}_{\mathcal{R}_2\mathcal{D}_1} \leq N_2^{'}. \qquad (5.22)
\end{aligned}
$$

The first eight inequality constraints are due to a Z channel[1] formed between

[1] Z channel is a special case of X channel where there is no message or channel from

any two transmitters and any two receivers as explained in [100]; and the last four inequality constraints are derived since the total degrees of freedom for a multi-antenna transmitter can not exceed the number of transmit antennas. Each of the first eight inequalities corresponds to one of the parameters $\mu'_{\mathcal{S}_i\mathcal{R}_j}$ or $\mu'_{\mathcal{R}_i\mathcal{D}_j}$, $\forall i,j \in \{1,2\}$ in the sense that if $\mu'_{\mathcal{S}_i\mathcal{R}_j}$ or $\mu'_{\mathcal{R}_i\mathcal{D}_j}$ is zero, the corresponding inequality is removed from the set of constraints. Besides that for $\forall i,j \in \{1,2\}$, we define:

$$\mu'_{ij} = \min\{\mu'_{\mathcal{S}_i\mathcal{R}_j}, \mu'_{\mathcal{R}_j\mathcal{D}_i}\} \tag{5.23}$$

$$\zeta_{ij} = \min\{\zeta_{\mathcal{S}_i\mathcal{R}_j}, \zeta_{\mathcal{R}_j\mathcal{D}_i}\}. \tag{5.24}$$

After choosing μ'_{ij} for each case, μ_{ij}; $i,j = 1,2$ are computed as function of μ'_{ij}; $i,j = 1,2$ as will be explained later. We briefly categorize the design process into the following steps:

1. How to compute the auxiliary variables (M'_1, M'_2, N'_1, N'_2) as functions of (M_1, M_2, N_1, N_2)?
2. After choosing $\mu'_{ij}; i,j = 1,2$, how to compute μ_{ij}?
3. How to choose the transmit and receive filters?

Case I: $N_1 \geq N_2 \geq M_1 \geq M_2$

---

one of the transmitters to one of the receivers [112].

This is an irreducible case, where we have the following equalities:

$$M_1^{'} = M_1, \tag{5.25}$$

$$M_2^{'} = M_2 \tag{5.26}$$

$$N_1^{'} = N_1, \tag{5.27}$$

$$N_2^{'} = N_2. \tag{5.28}$$

$\mu_{\mathcal{S}_i\mathcal{R}_j}^{'}$ and $\mu_{\mathcal{R}_j\mathcal{D}_i}^{'}$, $\forall(i,j) \in \{1,2\}$ are chosen subjects to constraints (5.22). We also have that:

$$\mu_{\mathcal{S}_1\mathcal{R}_1} = \mu_{\mathcal{S}_1\mathcal{R}_1}^{'} \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_1}^{'}, \tag{5.29}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_1} = \mu_{\mathcal{S}_2\mathcal{R}_1}^{'} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_1} = \mu_{\mathcal{R}_2\mathcal{D}_1}^{'}, \tag{5.30}$$

$$\mu_{\mathcal{S}_1\mathcal{R}_2} = \mu_{\mathcal{S}_1\mathcal{R}_2}^{'} \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_2} = \mu_{\mathcal{R}_1\mathcal{D}_2}^{'}, \tag{5.31}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_2} = \mu_{\mathcal{S}_2\mathcal{R}_2}^{'} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_2} = \mu_{\mathcal{R}_1\mathcal{D}_2}^{'}. \tag{5.32}$$

Consequently, $\mu_{ij}^{'}$, $i,j = 1,2$, are found from (5.23). $\boldsymbol{T}_{\mathcal{S}_1}$, $\boldsymbol{T}_{\mathcal{S}_2}$, $\boldsymbol{R}_{\mathcal{D}_1}$ and $\boldsymbol{R}_{\mathcal{D}_2}$ are randomly chosen from $\mathbb{OC}^{M_1\times(\mu_{11}+\mu_{12})}$, $\mathbb{OC}^{M_2\times(\mu_{21}+\mu_{22})}$, $\mathbb{OC}^{M_1\times(\mu_{11}+\mu_{12})}$ and $\mathbb{OC}^{M_2\times(\mu_{21}+\mu_{22})}$, respectively. We also have the following equalities:

$$\zeta_{\mathcal{S}_1\mathcal{R}_1} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_1\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{5.33}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_1} = \mu_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_2\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{5.34}$$

$$\zeta_{\mathcal{S}_1\mathcal{R}_2} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_1\mathcal{D}_2} = \mu_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}, \tag{5.35}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_2} = \mu_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_2\mathcal{D}_2} = \mu_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}. \tag{5.36}$$

Considering the transmit and receive filters described above, we define:

$$\widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_1} = \boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_1}\boldsymbol{T}_{\mathcal{S}_1} \quad , \quad \widetilde{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_1} = \boldsymbol{R}^*_{\mathcal{D}_1}\boldsymbol{H}_{\mathcal{R}_1\mathcal{D}_1}, \tag{5.37}$$

$$\widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_2} = \boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2}\boldsymbol{T}_{\mathcal{S}_1} \quad , \quad \widetilde{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_2} = \boldsymbol{R}^*_{\mathcal{D}_2}\boldsymbol{H}_{\mathcal{R}_1\mathcal{D}_2}, \tag{5.38}$$

$$\widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_1} = \boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_1}\boldsymbol{T}_{\mathcal{S}_2} \quad , \quad \widetilde{\boldsymbol{H}}_{\mathcal{R}_2\mathcal{D}_1} = \boldsymbol{R}^*_{\mathcal{D}_1}\boldsymbol{H}_{\mathcal{R}_2\mathcal{D}_1}, \tag{5.39}$$

$$\widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_2} = \boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2}\boldsymbol{T}_{\mathcal{S}_2} \quad , \quad \widetilde{\boldsymbol{H}}_{\mathcal{R}_2\mathcal{D}_2} = \boldsymbol{R}^*_{\mathcal{D}_2}\boldsymbol{H}_{\mathcal{R}_2\mathcal{D}_2}. \tag{5.40}$$

Therefore, $\boldsymbol{R}_{\mathcal{S}_i\mathcal{R}_j} \in \mathbb{OC}^{N_i\times(N_i-\zeta_{\mathcal{S}_i\mathcal{R}_j})}$ and $\boldsymbol{T}_{\mathcal{R}_j\mathcal{D}_i} \in \mathbb{OC}^{M_i\times(M_i-\zeta_{\mathcal{R}_j\mathcal{D}_i})}$ are chosen such that (Note that $\boldsymbol{A}\perp\boldsymbol{B}$ means that each column of matrix $\boldsymbol{A}$ is orthogonal to all columns of matrix $\boldsymbol{B}$):

$$\boldsymbol{R}_{\mathcal{S}_1\mathcal{R}_1}\perp\widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_1} \quad , \quad \boldsymbol{T}_{\mathcal{R}_1\mathcal{D}_1}\perp\widetilde{\boldsymbol{H}}^*_{\mathcal{R}_2\mathcal{D}_1} \tag{5.41}$$

$$\boldsymbol{R}_{\mathcal{S}_2\mathcal{R}_1}\perp\widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_1} \quad , \quad \boldsymbol{T}_{\mathcal{R}_2\mathcal{D}_1}\perp\widetilde{\boldsymbol{H}}^*_{\mathcal{R}_1\mathcal{D}_1}, \tag{5.42}$$

$$\boldsymbol{R}_{\mathcal{S}_1\mathcal{R}_2}\perp\widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \boldsymbol{T}_{\mathcal{R}_1\mathcal{D}_2}\perp\widetilde{\boldsymbol{H}}^*_{\mathcal{R}_2\mathcal{D}_2}, \tag{5.43}$$

$$\boldsymbol{R}_{\mathcal{S}_2\mathcal{R}_2}\perp\widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_2} \quad , \quad \boldsymbol{T}_{\mathcal{R}_2\mathcal{D}_2}\perp\widetilde{\boldsymbol{H}}^*_{\mathcal{R}_1\mathcal{D}_2}. \tag{5.44}$$

Considering all transmit and receive filters, the total equivalent channel matrices and the equivalent noise vectors are described by the following equations:

$$\widehat{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_1} = \boldsymbol{R}^*_{\mathcal{S}_1\mathcal{R}_1}\widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_1} \quad , \quad \widehat{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_1} = \widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_1}\boldsymbol{T}_{\mathcal{R}_1\mathcal{D}_1}, \tag{5.45}$$

$$\widehat{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_2} = \boldsymbol{R}^*_{\mathcal{S}_1\mathcal{R}_2}\widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_2} \quad , \quad \widehat{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_1} = \widetilde{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_2}\boldsymbol{T}_{\mathcal{R}_1\mathcal{D}_2}, \tag{5.46}$$

$$\widehat{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_1} = \boldsymbol{R}^*_{\mathcal{S}_2\mathcal{R}_1}\widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_1} \quad , \quad \widehat{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_1} = \widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_1}\boldsymbol{T}_{\mathcal{R}_2\mathcal{D}_1}, \tag{5.47}$$

$$\widehat{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_2} = \boldsymbol{R}^*_{\mathcal{S}_2\mathcal{R}_2}\widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \widehat{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_1} = \widetilde{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_2}\boldsymbol{T}_{\mathcal{R}_2\mathcal{D}_2}. \tag{5.48}$$

$$\boldsymbol{z}'_{\mathcal{R}_1} = \boldsymbol{R}^*_{\mathcal{S}_1\mathcal{R}_1}\boldsymbol{z}_{\mathcal{R}_1}, \tag{5.49}$$

$$\boldsymbol{z}''_{\mathcal{R}_1} = \boldsymbol{R}^*_{\mathcal{S}_2\mathcal{R}_1}\boldsymbol{z}_{\mathcal{R}_1}, \tag{5.50}$$

$$\boldsymbol{z}'_{\mathcal{R}_2} = \boldsymbol{R}^*_{\mathcal{S}_1\mathcal{R}_2}\boldsymbol{z}_{\mathcal{R}_2}, \tag{5.51}$$

$$\boldsymbol{z}''_{\mathcal{R}_2} = \boldsymbol{R}^*_{\mathcal{S}_2\mathcal{R}_2}\boldsymbol{z}_{\mathcal{R}_2}. \tag{5.52}$$

The first parallel MIMO relay channel viewed from the source-destination pair $(\mathcal{S}_1, \mathcal{D}_1)$ is represented by the following equations:

$$(\mathcal{S}_1, \mathcal{D}_1) \rightarrow \begin{cases} \boldsymbol{y}'_{\mathcal{R}_1} = \widehat{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_1}\boldsymbol{x}'_{\mathcal{S}_1} + \boldsymbol{z}'_{\mathcal{R}_1}, \\ \boldsymbol{y}'_{\mathcal{R}_2} = \widehat{\boldsymbol{H}}_{\mathcal{S}_1\mathcal{R}_2}\boldsymbol{x}''_{\mathcal{S}_1} + \boldsymbol{z}'_{\mathcal{R}_2}, \\ \boldsymbol{y}'_{\mathcal{D}_1} = \widehat{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_1}\boldsymbol{x}'_{\mathcal{R}_1} + \widehat{\boldsymbol{H}}_{\mathcal{R}_2\mathcal{D}_1}\boldsymbol{x}'_{\mathcal{R}_2} + \boldsymbol{z}'_{\mathcal{D}_1}, \end{cases} \tag{5.53}$$

$$(\mathcal{S}_2, \mathcal{D}_2) \rightarrow \begin{cases} \boldsymbol{y}''_{\mathcal{R}_1} = \widehat{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_1}\boldsymbol{x}'_{\mathcal{S}_2} + \boldsymbol{z}''_{\mathcal{R}_1}, \\ \boldsymbol{y}''_{\mathcal{R}_2} = \widehat{\boldsymbol{H}}_{\mathcal{S}_2\mathcal{R}_2}\boldsymbol{x}''_{\mathcal{S}_2} + \boldsymbol{z}''_{\mathcal{R}_2}, \\ \boldsymbol{y}'_{\mathcal{D}_2} = \widehat{\boldsymbol{H}}_{\mathcal{R}_1\mathcal{D}_2}\boldsymbol{x}'_{\mathcal{R}_1} + \widehat{\boldsymbol{H}}_{\mathcal{R}_2\mathcal{D}_2}\boldsymbol{x}'_{\mathcal{R}_2} + \boldsymbol{z}'_{\mathcal{D}_2}, \end{cases} \tag{5.54}$$

where

$$\begin{cases} \boldsymbol{z}'_{\mathcal{D}_1} = \boldsymbol{R}^*_{\mathcal{D}_1}\boldsymbol{z}_{\mathcal{D}_1}, \\ \boldsymbol{z}'_{\mathcal{D}_2} = \boldsymbol{R}^*_{\mathcal{D}_2}\boldsymbol{z}_{\mathcal{D}_2}. \end{cases} \tag{5.55}$$

and $\boldsymbol{x}'_{\mathcal{S}_i} \in \mathbb{C}^{\mu_{i1}\times 1}$, $\boldsymbol{x}''_{\mathcal{S}_i} \in \mathbb{C}^{\mu_{i2}\times 1}$, $\boldsymbol{x}'_{\mathcal{R}_i} \in \mathbb{C}^{\mu_{1i}\times 1}$ and $\boldsymbol{x}''_{\mathcal{R}_i} \in \mathbb{C}^{\mu_{2i}\times 1}$ are the data vectors transmitted over $\mathcal{S}_i \rightarrow \mathcal{R}_1$, $\mathcal{S}_i \rightarrow \mathcal{R}_2$, $\mathcal{R}_i \rightarrow \mathcal{D}_1$ and $\mathcal{R}_i \rightarrow \mathcal{D}_2$ links, respectively. The decomposition schemes proposed in section 5.3 simplify the signaling design and the performance evaluation for the X channels. However, such decomposition schemes deteriorate the performance of the system

because: (i) filters at the relay nodes are chosen such that the interference terms are forced to be zero, while the statistical properties of the interference should be exploited to design these filters, (ii) source transmit filters and destination receive filters are chosen randomly, while the gain of the channel matrices in the different directions should be considered.

5.4 Extension to Larger Networks

In the previous section, we considered a MIMO interference relay network which consists of only two sources, two relays and two destination nodes. Orthogonal beamforming was used at the source and destination nodes to multiplex and demultiplex the data streams respectively, while zero-forcing beamforming was used at the relay nodes to cancel the cross-link interferences. In the rest of this chapter, we extend that network model to an arbitrary number of sources, relays and destination nodes and use zero-forcing beamforming at the source and destination nodes and apply orthogonal beamforming at the relay nodes. In fact we switch the beamforming technique at the relay nodes with the one at the source and destination nodes.

The extended MIRN consists of L source nodes (denoted by $\mathcal{S}_1, \ldots, \mathcal{S}_L$), L destination nodes (denoted by $\mathcal{D}_1, \ldots, \mathcal{D}_L$) all paired together and equipped with M antennas, and K N-antenna $(K, N \geq M, KN > LM, N \leq LM)$ relays $\mathcal{R}_1, \mathcal{R}_2, ..., \mathcal{R}_K$ distributed between the source and destination nodes (see Figure 5.3). The number of antennas at each node is carefully chosen such that we do not use the null space of any link to multiplex the data

streams (as Case I in section 5.3).

Similar to the previous chapter, we consider *half-duplex* amplify-and-forward (AF) [53] relaying where data transmission from source to destination requires two channel hops and two non-overlapping time slots. During the first time slot, the source nodes transmit several data streams to the relays which are silent. We also assume that the destination nodes do not receive any signal directly from the source nodes because of large-scale fading effects. During the second time slot, while the source nodes are idle, a selected set of relays amplify and forward their received signal from the source to the destination nodes. All channels are considered frequency-flat block fading with

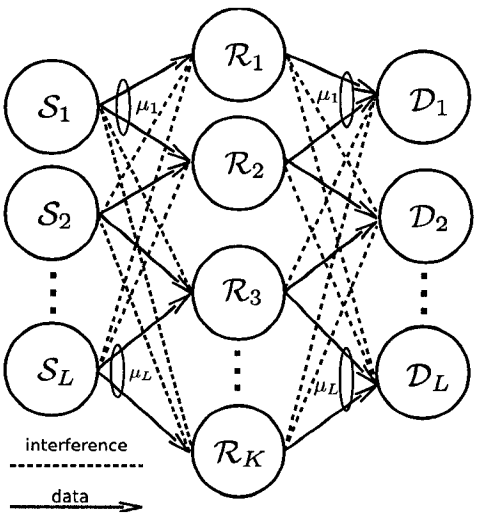

Figure 5.3 Block diagram of a two-hop MIMO interference relay nework.

independent realization across blocks and the duration of each block length is assumed to be equal to multiple time slots. During the first time slot,

$\mathcal{S}_i, 1 \leq i \leq L$, transmits a multiplexed data vector $\boldsymbol{x}_{\mathcal{S}_i} \in \mathbb{C}^{M\times 1}$ to *at most* μ_i different relays over the first hop where $\mu_i \in \mathbb{Z}^+$ is the spatial degrees of freedom available to the source node $\mathcal{S}_i$, such that a relay may be assigned with zero, one or even multiple data streams. For presentation simplicity, we write $\mathfrak{R}_i = \{\mathscr{R}_{i,1}, \ldots, \mathscr{R}_{i,\mu_i}\}$ to represent the set of selected relays associated with the source node $\mathcal{S}_i$. The source transmit vector can be written as:

$$\boldsymbol{x}_{\mathcal{S}_i} = \boldsymbol{T}_{\mathcal{S}_i}\boldsymbol{s}_i \tag{5.56}$$

where $\boldsymbol{s}_i \in \mathbb{C}^{\mu_i\times 1}$ is the data vector transmitted by $\mathcal{S}_i$, $\forall i \in \{1, 2, \ldots, L\}$ and $\boldsymbol{T}_{\mathcal{S}_i} = [\boldsymbol{\tau}_{\mathcal{S}_i,1}, \boldsymbol{\tau}_{\mathcal{S}_i,2}, \ldots, \boldsymbol{\tau}_{\mathcal{S}_i,\mu_i}] \in \mathbb{C}^{M\times\mu_i}$ is the spatial transmit filter designed to cancel the interference. We impose the following long-term power constraint on the source transmit vector:

$$\sum_{i=1}^{L} Tr\Big(\mathbb{E}\{\boldsymbol{x}_{\mathcal{S}_i}\boldsymbol{x}_{\mathcal{S}_i}^*\}\Big) \leq P_{\mathcal{S}}. \tag{5.57}$$

Suppose that $\mathfrak{S}_k$ denotes the set of source nodes that select relay $\mathcal{R}_k$ to convey part of their data streams. We can write the received vector at relay $\mathcal{R}_k$ as:

$$\begin{aligned}\boldsymbol{y}_{\mathcal{R}_k} &= \sum_{i=1}^{L} \boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k}\boldsymbol{x}_{\mathcal{S}_i} + \boldsymbol{z}_{\mathcal{R}_k} &(5.58)\\ &= \underbrace{\sum_{\forall\mathcal{S}_i\in\mathfrak{S}_k} \boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k}\boldsymbol{x}_{\mathcal{S}_i}}_{(Signal\,+\,Interference)} + \underbrace{\sum_{\forall\mathcal{S}_i\notin\mathfrak{S}_k} \boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k}\boldsymbol{x}_{\mathcal{S}_i}}_{(Interference)} + \boldsymbol{z}_{\mathcal{R}_k} &(5.59)\end{aligned}$$

where $\boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k} \in \mathbb{C}^{N\times M}$ and $\boldsymbol{z}_{\mathcal{R}_k} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N) \in \mathbb{C}^{N\times 1}$ denote the $\mathcal{S}_i \rightarrow \mathcal{R}_k$

channel matrix and the white zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise vector at relay $\mathcal{R}_k$, respectively. Relay $\mathcal{R}_k$ processes its received vector $\boldsymbol{y}_{\mathcal{R}_k}$ linearly by a matrix $\boldsymbol{G}_{\mathcal{R}_k} \in \mathbb{C}^{N\times N}$ and forwards it to the destination nodes. $\boldsymbol{G}_{\mathcal{R}_k}$ may be written as

$$\boldsymbol{G}_{\mathcal{R}_k} = \boldsymbol{V}_{\mathcal{R}_k}\boldsymbol{F}_{\mathcal{R}_k}\boldsymbol{U}^*_{\mathcal{R}_k} \tag{5.60}$$

where $\boldsymbol{V}_{\mathcal{R}_k} \in \mathbb{C}^{N\times\sum_i \mu_{ik}}$, $\boldsymbol{F}_{\mathcal{R}_k} \in \mathbb{C}^{\sum_i \mu_{ik}\times\sum_i \mu_{ik}}$ and $\boldsymbol{U}_{\mathcal{R}_k} \in \mathbb{C}^{N\times\sum_i \mu_{ik}}$ denote respectively the receive beamforming, gain and transmit beamforming matrices corresponding to relay $\mathcal{R}_k$[2]. The aggregate power constraint on the transmit vectors from the relay set $\mathfrak{R} = \bigcup_{i=1}^{L} \mathfrak{R}_i$ is[3]:

$$\sum_{\forall \mathcal{R}_k \in \mathfrak{R}} Tr\Big(\mathbb{E}\{\boldsymbol{x}_{\mathcal{R}_k}\boldsymbol{x}^*_{\mathcal{R}_k}\}\Big) = \sum_{\forall \mathcal{R}_k \in \mathfrak{R}} Tr\Big(\mathbb{E}\{(\boldsymbol{G}_{\mathcal{R}_k}\boldsymbol{y}_{\mathcal{R}_k})(\boldsymbol{G}_{\mathcal{R}_k}\boldsymbol{y}_{\mathcal{R}_k})^*\}\Big) \leq P_{\mathcal{R}}. \tag{5.61}$$

The transmit vectors from all selected relays ($\forall \mathcal{R}_k \in \mathfrak{R}$) then cross the $\mathcal{R}_k \rightarrow \mathcal{D}_i$ channels (which is denoted by $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}_i} \in \mathbb{C}^{M\times N}$) and add up with a white ZMCSCG noise vector $\boldsymbol{z}_{\mathcal{D}_i} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$ at destination $\mathcal{D}_i$. Assuming perfect synchronization between relays, we have the following received vector

[2] μ_{ik} denotes the number of data streams sent from $\mathcal{S}_i$ to $\mathcal{D}_i$ via $\mathcal{R}_k$. $\mu_i = \sum_k \mu_{ik}$ denotes the number of data streams sent by $\mathcal{S}_i$ to $\mathcal{D}_i$

[3] $\mathfrak{R}_i = \{\mathscr{R}_{i,1}, \ldots, \mathscr{R}_{i,\mu_i},\}$ denotes the set of relays selected by source-destination pair $(\mathcal{S}_i, \mathcal{D}_i)$.

at $\mathcal{D}_i$:

$$\boldsymbol{y}_{\mathcal{D}_i} = \sum_{\forall \mathcal{R}_k \in \mathfrak{R}} \boldsymbol{H}_{\mathcal{R}_k \mathcal{D}_i} \boldsymbol{x}_{\mathcal{R}_k} + \boldsymbol{z}_{\mathcal{D}_i} \tag{5.62}$$

$$= \underbrace{\sum_{\forall \mathcal{R}_k \in \mathfrak{R}_i} \boldsymbol{H}_{\mathcal{R}_k \mathcal{D}_i} \boldsymbol{x}_{\mathcal{R}_k}}_{Signal + Interference} + \underbrace{\sum_{\forall \mathcal{R}_k \notin \mathfrak{R}_i} \boldsymbol{H}_{\mathcal{R}_k \mathcal{D}_i} \boldsymbol{x}_{\mathcal{R}_k}}_{Interference} + \boldsymbol{z}_{\mathcal{D}_i}. \tag{5.63}$$

$\mathcal{D}_i$ multiplies $\boldsymbol{y}_{\mathcal{D}_i}$ by a spatial receive filter $\boldsymbol{R}_{\mathcal{D}_i} = [\boldsymbol{\rho}^T_{\mathcal{D}_i,1}, \boldsymbol{\rho}^T_{\mathcal{D}_i,2}, \ldots, \boldsymbol{\rho}^T_{\mathcal{D}_i,\mu_i}]^T \in \mathbb{C}^{\mu_i \times M}$ to separate the interfering data streams:

$$\tilde{\boldsymbol{s}}_i = \boldsymbol{R}_{\mathcal{D}_i} \boldsymbol{y}_{\mathcal{D}_i}, \tag{5.64}$$

where $\tilde{\boldsymbol{s}}_i \in \mathbb{C}^{\mu_i}$ is the vector of estimated data streams corresponding to $(\mathcal{S}_1, \mathcal{D}_1)$.

Since there are L source-destination pairs and each pair has μ_i independent non-interfering sub-channels, we may write the sum capacity for a given $\boldsymbol{T}_{\mathcal{S}_i}$, $\boldsymbol{R}_{\mathcal{D}_i}$ and $\boldsymbol{G}_{\mathcal{R}_k}, \mathcal{R}_k \in \mathfrak{R}$ as the summation of the signal to interference plus noise ratio (SINR) of these sub-channls:

$$C_{sum} = \sum_{i=1}^{L} \sum_{m=1}^{\mu_i} (0.5) \log(1 + \eta_{i,m}) \tag{5.65}$$

where $\eta_{i,m}$ is the $m^{th}, m = 1, \ldots, \mu_i$, sub-channel gain of the $i^{th}, i = 1, \ldots, L$, source-destination pair and the (0.5) factor is due to the half-duplex con-

straint. The value of $\eta_{i,m}$ may be expressed as [74]

$$\eta_{i,m} = \frac{\eta_{i,m}^{'}\eta_{i,m}^{''}}{1+\eta_{i,m}^{'}+\eta_{i,m}^{''}} \tag{5.66}$$

where $\eta_{i,m}^{'}$ and $\eta_{i,m}^{''}$ are the SINR of the first and second-hop component of the m^{th} compound sub-channel of $(\mathcal{S}_i, \mathcal{D}_i)$. We have the following expression for $\eta_{i,m}^{'}$ and $\eta_{i,m}^{''}$:

$$\eta_{i,m}^{'} = \frac{\sigma_{\mathcal{S}_i,m}^2 |\boldsymbol{u}_{\mathscr{R}_{i,m}}^* \boldsymbol{H}_{\mathcal{S}_i\mathscr{R}_{i,m}} \boldsymbol{\tau}_{\mathcal{S}_i,m}|^2}{1+\sum\sum_{\forall(j,n)\neq(i,m)} \sigma_{\mathcal{S}_j,n}^2 |\boldsymbol{u}_{\mathscr{R}_{i,m}}^* \boldsymbol{H}_{\mathcal{S}_j\mathscr{R}_{i,m}} \boldsymbol{\tau}_{\mathcal{S}_j,n}|^2}, \tag{5.67}$$

$$\eta_{i,m}^{''} = \frac{\sigma_{\mathscr{R}_{i,m}}^2 |\boldsymbol{\rho}_{\mathcal{D}_i,m} \boldsymbol{H}_{\mathscr{R}_{i,m}\mathcal{D}_i} \boldsymbol{v}_{\mathscr{R}_{i,m}}|^2}{1+\sum\sum_{\forall(j,n)\neq(i,m)} \sigma_{\mathscr{R}_{j,n}}^2 |\boldsymbol{\rho}_{\mathcal{D}_i,m} \boldsymbol{H}_{\mathscr{R}_{j,n}\mathcal{D}_i} \boldsymbol{v}_{\mathscr{R}_{i,m}}|^2}, \tag{5.68}$$

where $\boldsymbol{v}_{\mathscr{R}_{i,m}}$ and $\boldsymbol{u}_{\mathscr{R}_{i,m}}$ are columns of $\boldsymbol{V}_{\mathscr{R}_{i,m}}$ and $\boldsymbol{U}_{\mathscr{R}_{i,m}}$ respectively; $\sigma_{\mathcal{S}_i,m}^2$ and $\sigma_{\mathscr{R}_{i,m}}^2$ satisfy the following power constraint

$$\sum_{i=1}^{L}\sum_{m=1}^{\mu_i} \sigma_{\mathcal{S}_i,m}^2 ||\boldsymbol{\tau}_{\mathcal{S}_i,m}||^2 \leq P_{\mathcal{S}}, \tag{5.69}$$

$$\sum_{i=1}^{L}\sum_{m=1}^{\mu_i} \sigma_{\mathscr{R}_{i,m}}^2 \leq P_{\mathcal{R}}. \tag{5.70}$$

Now, the capacity of the MIRN can be written as the following optimization problem:

$$C = \max_{\boldsymbol{T}_{\mathcal{S}_i}, \boldsymbol{R}_{\mathcal{D}_i}, \boldsymbol{G}_{\mathcal{R}_k}} C_{sum} \tag{5.71}$$

subject to the power constraints in (5.69) and (5.70). Determining the optimal $\boldsymbol{T}_{\mathcal{S}_i}$, $\boldsymbol{R}_{\mathcal{D}_i}$ and $\boldsymbol{G}_{\mathcal{R}_k}$ is difficult in practice, especially for large K. In the sequel, we will discuss the value of degrees of freedom μ_i, $i = 1, \ldots, L$ allocated to each source node, the choice of selected relays by the source nodes $\mathfrak{R} = \{\mathscr{R}_{1,1}, \ldots, \mathscr{R}_{1,\mu_1}, \ldots, \mathscr{R}_{L,1}, \ldots, \mathscr{R}_{L,\mu_L}\}$, the spatial transmit and receive filters $\boldsymbol{T}_{\mathcal{S}_i}$ and $\boldsymbol{R}_{\mathcal{D}_i}; i = 1, \ldots, L$, at the source and destination nodes respectively and the relay gain matrices $\boldsymbol{G}_{\mathcal{R}_k}; \forall \mathcal{R}_k \in \mathfrak{R}$.

5.5 Problem Formulation

Suppose that $r_{ij}(\mathsf{SNR})$ denotes the rate of transmission from $\mathcal{S}_i$ to $\mathcal{D}_i$ through $\mathcal{R}_j$ ($\mathcal{S}_i \rightarrow \mathcal{R}_j \rightarrow \mathcal{D}_i$) such that the probability of error could be made arbitrarily small for this rate at signal to noise ratio SNR. For such rate function, the degrees of freedom μ_{ij} is defined as:

$$\mu_{ij} = \lim_{\mathsf{SNR}\rightarrow\infty} \frac{r_{ij}}{\log \mathsf{SNR}} \tag{5.72}$$

Now the degrees of freedom *region* of such networks (denoted by DOF) is defined as the matrix of degrees of freedom $[\mu_{ij}]$ corresponding to the simultaneously achievable rate matrices $[R_{ij}(\mathsf{SNR})]$ with the capacity region $C(\mathsf{SNR})$ [93]. Having $\mathcal{A} = \{1, \ldots, L\} \times \{1, \ldots, K\}$, the degrees of freedom region (DOF) is defined by:

We may write the degrees-of-freedom maximization as the following linear

$$DOF = \Bigg\{ [\mu_{ij}] \in \mathbb{Z}^{KL} : \forall [w_{ij}] \in \mathbb{Z}^{KL}, \sum_{(i,j)\in\mathcal{A}} w_{ij}\mu_{ij} \leq \lim \sup_{\mathsf{SNR}\to\infty} \Big[\sup_{[R_{ij}]\in C(\mathsf{SNR})} \sum_{(i,j)\in\mathcal{A}} \frac{w_{ij}R_{ij}}{\mathsf{SNR}} \Big] \Bigg\} \quad (5.73)$$

programming problem:

$$\max \sum_{i,j} \mu_{ij} \quad (5.74)$$

subject to the following inequality constraints for $\forall (i,j) \in \mathcal{A}$:

$$\begin{cases} \sum_{k=1}^{K} \mu_{kj} + \sum_{k=1}^{K} \mu_{ik} - \mu_{ij} \leq N, \\ \sum_{k=1}^{K} \mu_{ik} \leq M. \end{cases} \quad (5.75)$$

The first set of KL inequality constraints are due to a MAC formed at any receiving node as explained in [100] and the last L set of inequality constraints are derived since the total degrees of freedom of a multi-antenna transmitter can not exceed the number of transmit antennas. Now we have an integer linear programming (ILP) problem . To solve such problems, we may *relax* the integer constraint which will result in a lower bound to the optimal value of ILP. Another approach is the "*branch-and-bound*" algorithm that we follow to solve such problems. This algorithm works by *recursively* decomposing the problem into smaller ILP and use the lower bounds from LP relaxation to solve the sub-problems [113]. In the following, we assume that each source node is aware of the feasible region of the ILP in (5.74).

5.6 Sub-Optimum Signaling and Interference Alignment

In this section we mention the three challenges ahead in the calculation of transmit and receive filters, gain matrices and the values of degrees of freedom. Then we will present the proposed algorithm to address these issues.

1. To achieve the maximum total degrees of freedom of the MIMO X networks, every transmitting node requires global channel knowledge of the network. In the case of MIMO interference relay networks, this requirement enforces all transmitting nodes to have full knowledge of the channels of all other nodes in the first and second hop. As the number of relays increases, it will become much more troublesome for those nodes to acquire such channel information.

2. The degrees-of-freedom region of MIMO interference and MIMO X networks are achieved through interference alignment which requires long channel extensions in time or frequency to generate enough space dimensions for the received signal at the destination nodes. The length of channel extension increases as the number of transmit/receive pair or the number of constraints as in (5.75) the increases. However this extension of channel may not be always feasible (fast fading or narrow-band channels).

3. The choice of $\mu_{ij}, \forall(i,j) \in \mathcal{A}$ is based solely on an optimization problem which does not take into account the channel conditions. On the other hand, relay selection and interference alignment depend entirely on the

channel state. Therefore, we need to *jointly* optimize both.

In order to address all these issues, we pursue the following procedure. First, we randomly order the source-destination pairs from 1 to L and pick the first pair, namely $(\mathcal{S}_1, \mathcal{D}_1)$. Then, $(\mathcal{S}_1, \mathcal{D}_1)$ performs semi-orthogonal relay selection and beamforming as explained in the previous chapter. In this algorithm, the source and the destination perform zero-forcing (channel inversion) to multiplex and recover the data streams to/from the relay nodes respectively, while the set of beamforming direction from the source to the relays and from the relays to the destination form a *semi-orthogonal* set [13].

The selection and beamforming algorithm by $(\mathcal{S}_1, \mathcal{D}_1)$ divides the relay nodes into two sets: those which are selected by $(\mathcal{S}_1, \mathcal{D}_1)$ pair (denoted by $\mathfrak{R}_1$) and those which are not (denoted by $\mathfrak{R} - \mathfrak{R}_1$). Hence, $(\mathcal{S}_1, \mathcal{D}_1)$ pair generates a signal sub-space and an interference sub-space at $\mathfrak{R}_1$ and $\mathfrak{R}-\mathfrak{R}_1$, respectively. Beamforming and selection by $(\mathcal{S}_1, \mathcal{D}_1)$ also generates a signal sub-space at $\mathcal{D}_1$ and an interference sub-space at $\mathcal{D} - \mathcal{D}_1$ where $\mathcal{D}$ denotes the set of all destination nodes.

In the second step, the second randomly selected pair (denoted by $(\mathcal{S}_2, \mathcal{D}_2)$) performs selection and beamforming based on the knowledge of the first pair's transmit and receive filter matrices as well as gain matrices of $\mathfrak{R}_1$. Suppose that the set of selected relays by this pair is denoted by $\mathfrak{R}_2$. For any choice of the relays and the beamforming directions, $(\mathcal{S}_2, \mathcal{D}_2)$ generates a signal sub-space at $\mathfrak{R}_2$ and $\mathcal{D}_2$ while it generates an interference sub-space at $\mathfrak{R} - \mathfrak{R}_2$ and $\mathcal{D} - \mathcal{D}_2$. Selection and beamforming by $(\mathcal{S}_2, \mathcal{D}_2)$ must be such that the

beamforming directions must not only be semi-orthogonal to each other but also semi-orthogonal to the signal space generated by $(\mathcal{S}_1, \mathcal{D}_1)$ (or in other words be aligned in the interference sub-space generated by $(\mathcal{S}_1, \mathcal{D}_1)$).

Therefore, the relays and beamforming directions are selected such that they cast aligning interference at undesignated receivers and casts semi-orthogonal projection at designated receivers. The next steps follow from the same rule: "Cast aligning interference at the interference sub-space of undesired receivers and cast orthogonal projection to the signal sub-space of desired receivers".

Here's the detailed steps of this algorithm: Let the singular value decomposition of relay $\mathcal{R}_k$'s backward and forward channel be:

$$\boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k} = \boldsymbol{U}_{\mathcal{S}_i\mathcal{R}_k}\Lambda^{1/2}_{\mathcal{S}_i\mathcal{R}_k}\boldsymbol{V}^*_{\mathcal{S}_i\mathcal{R}_k}, \tag{5.76}$$

$$\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}_i} = \boldsymbol{U}_{\mathcal{R}_k\mathcal{D}_i}\Lambda^{1/2}_{\mathcal{R}_k\mathcal{D}_i}\boldsymbol{V}^*_{\mathcal{R}_k\mathcal{D}_i}, \tag{5.77}$$

where $\Lambda^{1/2}_{\mathcal{S}_i\mathcal{R}_k} \in \mathbb{C}^{N\times M}$ and $\Lambda^{1/2}_{\mathcal{R}_k\mathcal{D}_i} \in \mathbb{C}^{M\times N}$ are diagonal matrices comprising the eigenvalues of $\boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}_i}$ respectively, and $\boldsymbol{U}_{\mathcal{S}_i\mathcal{R}_k} \in \mathbb{C}^{N\times N}$, $\boldsymbol{U}_{\mathcal{R}_k\mathcal{D}_i} \in \mathbb{C}^{M\times M}$, $\boldsymbol{V}_{\mathcal{S}_i\mathcal{R}_k} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{V}_{\mathcal{R}_k\mathcal{D}_i} \in \mathbb{C}^{N\times N}$ are unitary matrices of eigenvectors.

Step 1: The source-destination pairs are randomly ordered. We assume that the random ordering of source nodes is $(\mathcal{S}_1, \mathcal{D}_1), (\mathcal{S}_2, \mathcal{D}_2), \ldots, (\mathcal{S}_L, \mathcal{D}_L)$. $(\mathcal{S}_1, \mathcal{D}_1)$ solves the ILP in 5.74 and among the set of solutions picks the largest μ_1. $(\mathcal{S}_1, \mathcal{D}_1)$ then performs beamforming and relay selection according to semi-orthogonal relay selection and beamfroming algorithm as described in

chapter 4. At the same time, this node informs the other source nodes of its beamforming matrix and its selected relays. Here we summarize the selection algorithm by $(\mathcal{S}_1, \mathcal{D}_1)$:

1. *Start-up*: For $(\mathcal{S}_1, \mathcal{D}_1)$ pair, we define a set of *modes* consisting of three indices indicating the index of the relay, index of the backward eigenmode and the index of the forward eigenmode. The selected set (search domain) is described by:

$$\mathcal{T}_{1,0} = \left\{ (\mathcal{R}_k, p, q) \middle| \lambda_{\mathcal{S}_1\mathcal{R}_k}(p) \geq \lambda_0; \lambda_{\mathcal{R}_k\mathcal{D}_1}(q) \geq \lambda_0 \right\} \tag{5.78}$$

where $\lambda_{\mathcal{S}_1\mathcal{R}_k}(p)$ and $\lambda_{\mathcal{R}_k\mathcal{D}_1}(q)$ are the p^{th} and q^{th} singular values of $\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_k}\boldsymbol{H}^*_{\mathcal{S}_1\mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}_1}\boldsymbol{H}^*_{\mathcal{R}_k\mathcal{D}_1}$, respectively, and λ_0 is a threshold value. The goal of this threshold is to limit the search space for semi-orthogonal eigenmodes and, most importantly, to constrain the amount of CSI feedback from the relays to the source. The proposed algorithm works as follows. The source chooses the mode with the largest effective channel gain. The metric for the effective channel gain of a mode is the ratio of product and sum of backward and forward channel gains of that mode which has been previously proposed in [71] for single-antenna relays:

$$(\mathscr{R}_{1,1}, b_{1,1}, f_{1,1}) = \arg \max_{(\mathcal{R}_k,p,q)\in\mathcal{T}_{1,0}} \frac{\lambda_{\mathcal{S}_1\mathcal{R}_k}(p).\lambda_{\mathcal{R}_k\mathcal{D}_1}(q)}{\lambda_{\mathcal{S}_1\mathcal{R}_k}(p) + \lambda_{\mathcal{R}_k\mathcal{D}_1}(q)} \tag{5.79}$$

where $(\mathscr{R}_{1,1}, b_{1,1}, f_{1,1})$ refers to relay $\mathscr{R}_{1,1}$ receiving from its backward channel eigenmode $b_{1,1}$ and transmitting over its channel eigenmode

$f_{1,1}$. After selecting $(\mathscr{R}_{1,1}, b_{1,1}, f_{1,1})$, we eliminate this mode from the search domain $\mathcal{T}_{1,0}$ to obtain a new search domain $\mathcal{T}_{1,1}$:

$$\mathcal{T}_{1,1} = \mathcal{T}_{1,0} - \{(\mathscr{R}_{1,1}, b_{1,1}, f_{1,1})\}. \tag{5.80}$$

Besides that, we define an empty set $\mathcal{M}$ to denote the set of selected modes and add $(\mathscr{R}_{1,1}, b_{1,1}, f_{1,1})$ to this set:

$$\mathcal{M} = \mathcal{M} \cup \{(\mathscr{R}_{1,1}, b_{1,1}, f_{1,1})\}. \tag{5.81}$$

Besides that for $\forall(\mathcal{R}_k, i, j) \in \mathcal{T}_{1,1}$ we define:

$$\Gamma_{1,1}(i,j,k) = \max\left\{\gamma'_{1,1}(i,k), \gamma''_{1,1}(j,k)\right\} \tag{5.82}$$

where $\gamma_{1,m}(i,k)$ and $\gamma_{1,m}(j,k)$ are defined as:

$$\gamma'_{1,m}(i,k) = \varphi\Big(\boldsymbol{v}_{\mathcal{S}_1\mathscr{R}_{1,m}}(b_{1,m}), \boldsymbol{v}_{\mathcal{S}_1\mathcal{R}_k}(i)\Big) \tag{5.83}$$

$$\gamma''_{1,m}(j,k) = \varphi\Big(\boldsymbol{u}_{\mathscr{R}_{1,m}\mathcal{D}_1}(f_{1,m}), \boldsymbol{u}_{\mathcal{R}_k\mathcal{D}_1}(j)\Big) \tag{5.84}$$

and $\boldsymbol{v}_{\mathcal{S}_1\mathcal{R}_k}(i)$ and $\boldsymbol{u}_{\mathcal{R}_k\mathcal{D}_1}(j)$ are the i^{th} right and j^{th} left eigenvectors of $\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}_1}$, respectively, and $\varphi(\boldsymbol{a}, \boldsymbol{b})$ is defined in equation (4.18).

2. *Iterations*: For $2 \leq m \leq \mu_1$, the source performs the following steps:

$$(\mathscr{R}_{1,m}, b_{1,m}, f_{1,m}) = \arg \min_{(\mathcal{R}_k,i,j)\in\mathcal{T}_{1,m-1}} \Gamma_{1,m-1}(i,j,k) \tag{5.85}$$

$$\mathcal{T}_{1,m} = \mathcal{T}_{1,m-1} - \{(\mathscr{R}_{1,m}, b_{1,m}, f_{1,m})\} \tag{5.86}$$

$$\mathcal{M} = \mathcal{M} \cup \{(\mathscr{R}_{1,m}, b_{1,m}, f_{1,m})\}. \tag{5.87}$$

$$\Gamma_{1,m}(i,j,k) = \Gamma_{1,m}(i,j,k) + \max\left\{\gamma'_{1,m}(i), \gamma''_{1,m}(j)\right\},$$
$$\forall(\mathcal{R}_k, i, j) \in \mathcal{T}_{1,m}. \tag{5.88}$$

Now for the k^{th} selected mode by $(\mathcal{S}_1, \mathcal{D}_1)$, we may write the processing matrix as:

$$\boldsymbol{G}_{\mathscr{R}_{1,k}} = g_{1,k}\boldsymbol{v}_{\mathscr{R}_{1,k}\mathcal{D}_1}(f_{1,k})\boldsymbol{u}^*_{\mathcal{S}_1\mathscr{R}_{1,k}}(b_{1,k}) \tag{5.89}$$

where the coefficients $g_{1,k}, 1 \leq k \leq \mu_1$ satisfy the aggregate power constraint in (5.70). Similar to parallel MIMO relay channels introduced in the previous section, we may write the the effective channel matrix for the first-hop and second-hop of $(\mathcal{S}_1, \mathcal{D}_1)$ pair by:

$$\boldsymbol{\mathcal{H}}_{\mathcal{S}_1\mathfrak{R}_1} = \begin{bmatrix} \sqrt{\lambda_{\mathcal{S}_1\mathscr{R}_{1,1}}(b_{1,1})}\boldsymbol{v}^*_{\mathcal{S}_1\mathscr{R}_{1,1}}(b_{1,1}) \\ \sqrt{\lambda_{\mathcal{S}_1\mathscr{R}_{1,2}}(b_{1,2})}\boldsymbol{v}^*_{\mathcal{S}_1\mathscr{R}_{1,2}}(b_{1,2}) \\ \vdots \\ \sqrt{\lambda_{\mathcal{S}_1\mathscr{R}_{1,\mu_1}}(b_{1,\mu_1})}\boldsymbol{v}^*_{\mathcal{S}_1\mathscr{R}_{1,\mu_1}}(b_{1,\mu_1}) \end{bmatrix} \in \mathbb{C}^{M\times M}, \tag{5.90}$$

and

$$\boldsymbol{\mathcal{H}}_{\mathfrak{R}_1\mathcal{D}_1} = \begin{bmatrix} \sqrt{\lambda_{\mathscr{R}_{1,1}\mathcal{D}_1}(f_{1,1})}\boldsymbol{u}^T_{\mathscr{R}_{1,1}\mathcal{D}}(f_{1,1}) \\ \sqrt{\lambda_{\mathscr{R}_{1,2}\mathcal{D}_1}(f_{1,2})}\boldsymbol{u}^T_{\mathscr{R}_{1,2}\mathcal{D}}(f_{1,2}) \\ \vdots \\ \sqrt{\lambda_{\mathscr{R}_{1,\mu_1}\mathcal{D}_1}(f_{1,\mu_1})}\boldsymbol{u}^T_{\mathscr{R}_{1,\mu_1}\mathcal{D}}(f_{1,M}) \end{bmatrix}^T \in \mathbb{C}^{M\times M}. \tag{5.91}$$

Therefore, the transmit and receive filters deployed by $(\mathcal{S}_1, \mathcal{D}_1)$ are respectively pseudo-inverses of $\boldsymbol{\mathcal{H}}_{\mathcal{S}_1\mathfrak{R}_1}$ and $\boldsymbol{\mathcal{H}}_{\mathfrak{R}_1\mathcal{D}_1}$:

$$\boldsymbol{T}_{\mathcal{S}_1} = \boldsymbol{\mathcal{H}}^{\dagger}_{\mathcal{S}_1\mathfrak{R}_1}, \tag{5.92}$$

$$\boldsymbol{R}_{\mathcal{D}_1} = \boldsymbol{\mathcal{H}}^{\dagger}_{\mathfrak{R}_1\mathcal{D}_1}. \tag{5.93}$$

Step 2 to L: By inserting the values of μ_{jk}, $1 \leq j < i$, into (5.74), $(\mathcal{S}_i, \mathcal{D}_i)$ resolves (5.74) and searches the set of solution for the largest μ_i that maximizes (5.74). Having known the filter matrices of $(\mathcal{S}_1, \mathcal{D}_1)$,..., $(\mathcal{S}_{i-1}, \mathcal{D}_{i-1})$ and the set of selected relays $\mathfrak{R}_1$,..., $\mathfrak{R}_{i-1}$ and their processing matrices, $(\mathcal{S}_i, \mathcal{D}_i)$ finds the beamforming directions that are semi-orthogonal to the signal sub-space and are aligned with the interference subspace previous source-destination pairs.

1. *Initialization*: For $(\mathcal{S}_i, \mathcal{D}_i)$ pair, we define a set of *modes* consisting of three indices indicating the index of the relay, index of the backward eigenmode and the index of the forward eigenmode. The selected set

(search domain) is described by:

$$\mathcal{T}_{i,0} = \left\{ (\mathcal{R}_k, p, q) \middle| \lambda_{\mathcal{S}_i\mathcal{R}_k}(p) \geq \lambda_0; \lambda_{\mathcal{R}_k\mathcal{D}_i}(q) \geq \lambda_0 \right\} \tag{5.94}$$

where $\lambda_{\mathcal{S}_i\mathcal{R}_k}(p)$ and $\lambda_{\mathcal{R}_k\mathcal{D}_i}(q)$ are the p^{th} and q^{th} singular values of $\boldsymbol{H}_{\mathcal{S}_i\mathcal{R}_k}\boldsymbol{H}^*_{\mathcal{S}_i\mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k\mathcal{D}_i}\boldsymbol{H}^*_{\mathcal{R}_k\mathcal{D}_i}$, respectively, and λ_0 is a threshold value.

2. *Start-up*: The first selected mode is found from:

$$(\mathscr{R}_{i,1}, b_{i,1}, f_{i,1}) = \arg \min_{(\mathcal{R}_k,\alpha,\beta)\in\mathcal{T}_{i,0}} \Gamma_{i,0}(\alpha, \beta, k) \tag{5.95}$$

where $(\mathscr{R}_{i,1}, b_{i,1}, f_{i,1})$ refers to relay $\mathscr{R}_{i,1}$ receiving from its backward channel eigenmode $b_{i,1}$ and transmitting over its channel eigenmode $f_{i,1}$; and $\Gamma_{i,0}(\alpha, \beta, k)$ is defined as:

$$\Gamma_{i,0}(\alpha, \beta, k) = \max\left\{ \gamma'_{i,0}(\alpha, k), \gamma''_{i,0}(\beta, k) \right\} \tag{5.96}$$

where $\gamma'_{i,0}(\alpha, k)$ and $\gamma''_{i,0}(\beta, k)$ are defined as:

$$\gamma'_{i,0}(\alpha, k) = \sum\nolimits_{\forall(\mathscr{R}_{j,n},x,y)\in\mathcal{M}} \left(\boldsymbol{u}^*_{\mathscr{R}_{j,n}}(x)\boldsymbol{H}_{\mathcal{S}_i\mathscr{R}_{j,n}} \right) \left(\sqrt{\lambda_{\mathcal{S}_i\mathcal{R}_k}(\alpha)}\boldsymbol{v}_{\mathcal{S}_i\mathcal{R}_k}(\alpha) \right), \tag{5.97}$$

$$\gamma''_{i,0}(\beta, k) = \sum\nolimits_{\forall(\mathscr{R}_{j,n},x,y)\in\mathcal{M}} \left(\sqrt{\lambda_{\mathscr{R}_i\mathcal{D}_i}(\beta)}\boldsymbol{u}_{\mathcal{R}_k\mathcal{D}_i}(\beta) \right) \left(\boldsymbol{H}_{\mathscr{R}_{j,n}\mathcal{D}_i}\boldsymbol{v}^*_{\mathscr{R}_{j,n}}(y) \right). \tag{5.98}$$

In fact, $\gamma'_{i,0}(\alpha, k)$ and $\gamma''_{i,0}(\beta, k)$ represent the amount of projection of the candidate mode over previously selected modes over the first and second hop, respectively. After selecting $(\mathscr{R}_{i,1}, b_{i,1}, f_{i,1})$, we eliminate this mode from the search domain $\mathcal{T}_{i,0}$ and add to the set of elected

modes to obtain:

$$\mathcal{T}_{i,1} = \mathcal{T}_{i,0} - \{(\mathscr{R}_{i,1}, b_{i,1}, f_{i,1})\}, \tag{5.99}$$

$$\mathcal{M} = \mathcal{M} \cup (\mathscr{R}_{i,1}, b_{i,1}, f_{i,1}). \tag{5.100}$$

Besides that for $\forall(\mathcal{R}_k, \alpha, \beta) \in \mathcal{T}_{i,1}$ we define:

$$\Gamma_{i,1}(\alpha, \beta, k) = \max\left\{\gamma'_{i,1}(\alpha, k), \gamma''_{i,1}(\beta, k)\right\} \tag{5.101}$$

where $\gamma'_{i,m}(\alpha, k)$ and $\gamma''_{i,m}(\beta, k)$ are defined as:

$$\gamma'_{i,1}(\alpha, k) = \gamma'_{i,0}(\alpha, k) + (\boldsymbol{u}^*_{\mathscr{R}_{j,n}} \boldsymbol{H}_{\mathcal{S}_i \mathscr{R}_{j,n}})(\sqrt{\lambda_{\mathcal{S}_i \mathcal{R}_k}(\alpha)} \boldsymbol{v}_{\mathcal{S}_i \mathcal{R}_k}(\alpha)), \tag{5.102}$$

$$\gamma''_{i,1}(\beta, k) = \gamma''_{i,0}(\beta, k) + (\boldsymbol{v}^*_{\mathscr{R}_{j,n}} \boldsymbol{H}_{\mathscr{R}_{j,n} \mathcal{D}_i})(\sqrt{\lambda_{\mathscr{R}_i \mathcal{D}_i}(\beta)} \boldsymbol{u}_{\mathcal{R}_k \mathcal{D}_i}(\beta)). \tag{5.103}$$

and $\boldsymbol{v}_{\mathcal{S}_i \mathcal{R}_k}(\alpha)$ and $\boldsymbol{u}_{\mathcal{R}_k \mathcal{D}_i}(\beta)$ are the α^{th} right and β^{th} left eigenvectors of $\boldsymbol{H}_{\mathcal{S}_i \mathcal{R}_k}$ and $\boldsymbol{H}_{\mathcal{R}_k \mathcal{D}_i}$, respectively.

3. *Iterations*: For $2 \leq m \leq \mu_i$, the source performs the following steps:

$$(\mathscr{R}_{i,m}, b_{i,m}, f_{i,m}) = \arg \min_{(\mathcal{R}_k, \alpha, \beta) \in \mathcal{T}_{i,m-1}} \Gamma_{i,m-1}(\alpha, \beta, k) \tag{5.104}$$

$$\mathcal{T}_{i,m} = \mathcal{T}_{i,m-1} - \{(\mathscr{R}_{i,m}, b_{i,m}, f_{i,m})\} \tag{5.105}$$

$$\mathcal{M} = \mathcal{M} \cup \{(\mathscr{R}_{i,m}, b_{i,m}, f_{i,m})\}. \tag{5.106}$$

$$\Gamma_{i,m}(\alpha, \beta, k) = \max\left\{\gamma'_{i,m}(\alpha), \gamma''_{i,m}(\beta)\right\}, \quad \forall(\mathcal{R}_k, \alpha, \beta) \in \mathcal{T}_{i,m}. \tag{5.107}$$

Now for the k^{th} selected mode by $(\mathcal{S}_i, \mathcal{D}_i)$, we may write the processing matrix as:

$$\boldsymbol{G}_{\mathscr{R}_{i,k}} = g_{i,k}\boldsymbol{v}_{\mathscr{R}_{i,k}\mathcal{D}_i}(f_{i,k})\boldsymbol{u}^*_{\mathcal{S}_i\mathscr{R}_{i,k}}(b_{i,k}) \tag{5.108}$$

where the coefficients $g_{i,k}, 2 \leq i \leq L, 1 \leq k \leq \mu_i$, satisfy the aggregate power constraint in (5.70). Similarly to the previous chapter, we may write the the effective channel matrix from $\mathcal{S}_i$ to $\mathfrak{R}_i$ and from $\mathfrak{R}_i$ to $\mathcal{D}_i$ as the following:

$$\boldsymbol{\mathcal{H}}_{\mathcal{S}_i\mathfrak{R}_i} = \begin{bmatrix} \sqrt{\lambda_{\mathcal{S}_i\mathscr{R}_{i,1}}(b_{i,1})}\boldsymbol{v}^*_{\mathcal{S}_i\mathscr{R}_{i,1}}(b_{i,1}) \\ \sqrt{\lambda_{\mathcal{S}_i\mathscr{R}_{i,2}}(b_{i,2})}\boldsymbol{v}^*_{\mathcal{S}_i\mathscr{R}_{i,2}}(b_{i,2}) \\ \vdots \\ \sqrt{\lambda_{\mathcal{S}_i\mathscr{R}_{i,\mu_i}}(b_{i,\mu_i})}\boldsymbol{v}^*_{\mathcal{S}_i\mathscr{R}_{i,\mu_i}}(b_{i,\mu_i}) \end{bmatrix} \in \mathbb{C}^{M\times M}, \tag{5.109}$$

and

$$\boldsymbol{\mathcal{H}}_{\mathfrak{R}_i\mathcal{D}_i} = \begin{bmatrix} \sqrt{\lambda_{\mathscr{R}_{i,1}\mathcal{D}_i}(f_{i,1})}\boldsymbol{u}^T_{\mathscr{R}_{i,1}\mathcal{D}}(f_{i,1}) \\ \sqrt{\lambda_{\mathscr{R}_{i,2}\mathcal{D}_i}(f_{i,2})}\boldsymbol{u}^T_{\mathscr{R}_{i,2}\mathcal{D}}(f_{i,2}) \\ \vdots \\ \sqrt{\lambda_{\mathscr{R}_{i,\mu_i}\mathcal{D}_i}(f_{i,\mu_i})}\boldsymbol{u}^T_{\mathscr{R}_{i,\mu_i}\mathcal{D}}(f_{i,\mu_i}) \end{bmatrix}^T \in \mathbb{C}^{M\times M}. \tag{5.110}$$

Therefore, the transmit and receive filters deployed by $(\mathcal{S}_i, \mathcal{D}_i)$ are respectively pseudo-inverses of $\boldsymbol{\mathcal{H}}_{\mathcal{S}_1\mathfrak{R}_1}$ and $\boldsymbol{\mathcal{H}}_{\mathfrak{R}_1\mathcal{D}_1}$:

$$\boldsymbol{T}_{\mathcal{S}_i} = \boldsymbol{\mathcal{H}}^{\dagger}_{\mathcal{S}_i\mathfrak{R}_i}, \tag{5.111}$$

$$\boldsymbol{R}_{\mathcal{D}_i} = \boldsymbol{\mathcal{H}}^{\dagger}_{\mathfrak{R}_i\mathcal{D}_i}. \tag{5.112}$$

5.7 Simulation Results

In this section, numerical results for the proposed algorithm are presented. In Figure 5.4 we have plotted the ergodic capacity of the proposed algorithm (interference alignment) together with the capacity of upper bound, coherent relaying [11] and full cooperation. By full cooperation we mean that all the source and destination nodes are connected and perform semi-orthogonal relay selection and beamforming as described in chapter 4. The source and relay transmit power is assumed to be constrained to 10 dB. All channels are assumed to be zero-mean unit covariance Rayleigh fading channel while the effect of large-scale fading is neglected. We also set the value of the threshold equal to $\log K - \log\log K$.

We observe that although interference alignment offers lower degrees of freedom than coherent relaying, it has higher capacity. The reason is that by sacrificing some degrees of freedom, we may achieve higher SINR which compensate for the loss in the degrees of freedom. We can also see that if all the source and destination pair were fully cooperative, the rate of capacity growth would remain the same as interference alignment.

5.8 Conclusions

In this chapter, we considered the applications of MIMO X channel to the so-called "MIMO interference relay networks (MIRN)". The best known techniques which achieve the capacity of such channel required distributed

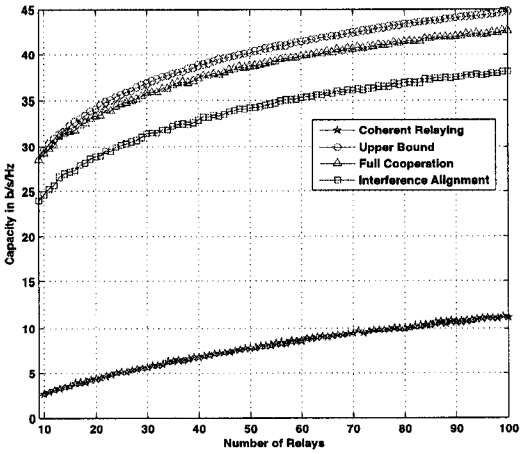

Figure 5.4 Ergodic capacity as a function of number of relays for $M = 3$, $N = 3$, $L = 3$ and $SNR = 10\ dB$.

interference cancellation in the asymptote of large number of relays or perfect cooperation among relay terminals. Based on the information-theoretical results that highlights the effect of *interference alignment* in enhancing the multiplexing gain of interference and X channels, we showed that cascading the first and second as two X channel is an alternative to the conventional techniques. Maximum degrees of freedom and interference alignment were discussed and simulation results confirmed the theory that for small to medium number of relays, decomposing MIRN into several parallel MIMO relay channels outperforms the conventional "distributed interference cancellation".

CHAPTER 6

CONCLUSIONS

6.1 Concluding Remarks

The main focus of this thesis is on "relay selection" and "beamforming" for two-hop MIMO (parallel or interference) relay channels. We show that two-hop interference cancellation technique outperforms conventional distributed interference cancellation for MIMO multi-hop networks for practical number of relay nodes.

In Chapter 2, we provided the necessary literature survey and provide definitions for the performance measures, i.e., ergodic and outage capacity, multiplexing gain, array gain and diversity gain for MIMO channels. We investigated the potential benefits of MIMO channels in terms of these gains compared to SISO channels.

In Chapter 3, we introduced the relay channels and two classes of relay nodes, namely FD and TDD. We showed that FD relay channels outperform direct link channels at all SNR's. However, superiority of TDD relay channels over direct link channels with no relaying depends strictly on $\mathcal{S} \rightarrow \mathcal{R}$, $\mathcal{R} \rightarrow \mathcal{D}$ and $\mathcal{S} \rightarrow \mathcal{D}$ channels. Since there are RF constraints for having FD relays, TDD relaying and in particular HD relaying are considered as potential candidates for relaying. For HD relay channels we investigated three different

transmission protocols together with AF and DF relay nodes. We showed that capacity performance of AF and DF relaying under these three protocols as function of $\mathsf{SNR}_{\mathcal{S}\to\mathcal{D}}$, $\mathsf{SNR}_{\mathcal{S}\to\mathcal{R}}$ and $\mathsf{SNR}_{\mathcal{R}\to\mathcal{D}}$. We observed that protocol I outperforms other protocols because of its multiplexing gain but has much more transceiver complexity than protocol II. As a rule of thumb for protocol II, we may say that DF relaying is beneficial when the relay node is close to the source node and AF relaying is helpful when the relay node is closer to the destination node than to the source node. DF relaying strategy enforces several constraints on the transmission rate for $\mathcal{S}\to\mathcal{R}$, $\mathcal{R}\to\mathcal{D}$ and $\mathcal{S}\to\mathcal{D}$ links. As the number of relay nodes increases, the number of these constraints grow exponentially. Besides that, there are non-trivial implementation issues associated with DF relaying which makes them less popular than HD-AF relaying. Therefore, we selected HD-AF relaying scheme in the sequel of this thesis.

In Chapter 4, we studied the MIMO parallel relay channels where data transmission from the source node to destination node is assisted by a large number of relay nodes which are located in the middle region between the source and destination. Two threshold-based, low-feedback, low-complexity relay selection algorithms namely, *eigenmode combining* and *antenna combining* were introduced based on semi-orthogonality among the backward and forward eigenmodes or antennas of the relays. Capacity scaling of relay selection is also discussed in Chapter 4. The main result of this chapter is that applying MIMO broadcast techniques (e.g. zero-forcing beamforming) to the first hop and MIMO multiple-access techniques (e.g. zero-forcing beamforming) at the second hop will outperform the conventional distributed interference cancel-

lation techniques. This is mainly due to capacity scaling of $\frac{M}{2}\log\log K+\mathcal{O}(1)$ compared to $\frac{M}{2}\log K+\mathcal{O}(1)$ for distributed interference cancellation techniques. In fact for small to medium number of relay nodes, the loss in array gain from $\log K$ to $\log\log K$ is compensated by the superiority of $\mathcal{O}(1)$ term in the selection algorithms than that of the distributed interference cancellation techniques.

In Chapter 5, we studied the MIMO interference relay network (MIRN) where multiple source-destination nodes communicate via multiple relay nodes which are located between the source and destination nodes. By introducing the concept of "interference alignment" and treating the first and second hop as two cascaded MIMO X channels, we jointly optimized the degree-of-freedom (DOF) and beamforming directions of the network. Simulation results showed that similarly to the previous chapter, two-hop interference cancellation outperforms distributed interference cancellation.

6.2 Suggestions for Future Works

Throughout this thesis, we assumed dedicated relay nodes which do not have any information of their own to transmit. On the other hand, the source and destination nodes do not assist any node as a relay. In a more general case, any node (whether it be source, relay or destination) may be a relay, destination or source for another node. Besides that, in ad hoc networks there are scenarios where there is a direct link from the source to the destination which makes relaying inefficient in terms of multiplexing gain. Addressing

this issue requires the application of random graph theory. A brief summary of the possible extension works is listed below.

1. Extension of the relay selection and beamforming algorithms to heterogeneous networks where the constraint on the location of the relays (to be in the middle region between the source and the destination) is relaxed,
2. Relay deployment in cellular networks and extension of MIRN to MIMO interfering cells,
3. Resource allocation and optimization for distributed MIMO ad hoc networks,
4. Cross layer design of an ad hoc network where every node may serve as a source, relay or destination for other nodes.

BIBLIOGRAPHY

[1] D. Gesbert, M. Shafi, D. shan Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 281–302, April 2003.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communication*, vol. 6, pp. 311–335, March 1998.

[3] V. Tarokh, N. Seshadri, and A. Calderblank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[4] E. Telatar, "Capacity of multiantenna Gaussian channels," *AT&T Bell Laboratories Technical Memorandum*, June 1995.

[5] R. Pabst, B. Walk, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukhrejee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. aghvami, D. Falconer, and G. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Communication Magazine*, vol. 42, no. 9, pp. 80–89, 2004.

[6] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—part I: System description," *IEEE Transactions on Communication*, vol. 51, no. 11, pp. 1927–1938, Nov 2003.

[7] ——, "User cooperation diversity—part II: Implementation aspects and performance analysis," *IEEE Transactions on Communication*, vol. 51, no. 11, pp. 1939–1948, November 2003.

[8] R. U. Nabar, H. Bölcskei, and F. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1099–1109, August 2004.

[9] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 29–43, January 2005.

[10] R. U. Nabar, O. Oyman, H. Bölcskei, and A. J. Paulraj, "Capacity scaling laws in MIMO wireless networks," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, October 2003, pp. 378–389.

[11] H. Bölcskei, R. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[12] O. Oyman and A. J. Paulraj, "Design and analysis of linear distributed MIMO relaying algorithms," *IEE Proceedings on Communications*, vol. 153, no. 4, pp. 565–572, August 2006.

[13] S. Oveis Gharan, A. Bayesteh, and A. K. Khandani, "Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay

network," *IEEE Transactions on Information Theory*, sumitted 2007. [Online]. Available: http://arxiv.org/abs/cs/0703151

[14] M. A. Torabi and J. F. Frigon, "Semi-orthogonal relay selection and beaforming for amplify-and-forward MIMO relay channels," in *Proceeding of IEEE Wireless Communications and Networking Conference (WCNC'08)*, Las Vegas, USA, 31 March - 3 April 2008.

[15] ——, "A decomposition approach to MIMO interference relay channels," in *Proceedings of IEEE GLobal Telecommunications Conference (Globecom'08)*, New Orleans/LA, USA, 30 November - 3 December 2008.

[16] A. Poon, R. W. Broaderson, and D. N. C. Tse, "Degrees of freedom in multiple-antenna channels: a signal space approach," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 523–536, February 2005.

[17] D. Tse and P. Viswanath, *Fundamentals of Wireless Commuications.* Cambridge University Press, 2005.

[18] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information theoretic and communication aspects," *IEEE Transactions on Information Theory*, vol. 44, pp. 2619–1692, October 1998.

[19] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[20] B. Hassibi and M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.

[21] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.

[22] H. ElGamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieves the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 968–985, June 2004.

[23] M. Jankiraman, *Space-time codes and MIMO systems.* Artech House Publishers, 2004.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Thoery.* Wiley, 1991.

[25] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side infromation," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, November 1997.

[26] S. K. Jayaweera and H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2697–2709, October 2003.

[27] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Transactions on Communication*, vol. 47, pp. 1458–1461, October 1999.

[28] P. A. Dighe, R. K. Mallik, and S. R. Jamur, "Analysis of transmit-receive diversity in rayleigh fading," in *Proceedings of IEEE Global Telecommunications Conference (Globecom'2001)*, San Antoni, TX, November 2001, pp. 1132–1136.

[29] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex wishart matrices and performance analysis of MIMO MRC systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 418–426, April 2003.

[30] J. B. Andersen, "Antenna arrays in mobile communications: gain, diversity, and channel capacity," *IEEE Antennas and Propagation Magazine*, pp. 12–16, April 2000.

[31] ——, "Array gain and capacity for known random channels with multiple element array at both ends," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2172–2178, November 2000.

[32] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, October 1998.

[33] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmanian beamforming for multiple-input multiple-output wireless system," *IEEE Transactions on Information Theory*, vol. 49, pp. 2735–2747, October 2003.

[34] K. K. Mukkavilli, A. Sabharval, E. Ekrip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antena systems," *IEEE*

Transactions on Information Theory, vol. 49, pp. 2562–2579, October 2003.

[35] D. J. Love, R. W. Heath Jr, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communication Magazine*, vol. 42, no. 10, pp. 54–59, October 2004.

[36] Z. W. S. Zhou and G. B. Giannakis, "Quantifying the power loss when transmit beamforming relies on finite-rate feedback," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1948–1957, 2005.

[37] X. Tang and Y. Hua, "Optimal design of non-regenrative MIMO wireless relays," *IEEE Transaction on Wireless Communications*, vol. 6, no. 4, pp. 1398–1407, April 2007.

[38] G. H. Golub and C. F. V. Loan, *Matrix computations*, 3rd ed. The John Hopkins University Press, 1996.

[39] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, April 1996.

[40] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communication Magazine*, vol. 42, no. 10, pp. 68–73, October 2004.

[41] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, University of California-Berkeley, June 1968.

[42] ——, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, no. 120-154, 1971.

[43] ——, "A survey of multiway channels in information theory: 1961–1976," *IEEE Transactions on Information Theory*, vol. IT-23, no. 1, pp. 1–37, January 1977.

[44] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. IT-25, pp. 572–584, September 1979.

[45] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, September 2005.

[46] M. Gastpar and M. Vetterli, "On the capacity of large gaussian relay networks," *IEEE Transactions on Information Theory*, vol. 51, pp. 765–779, 2005.

[47] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Transactions on Communications*, vol. 52, no. 9, pp. 1470–1476, September 2004.

[48] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.

[49] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: the relay case," *Proceedings of Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, 2002.

[50] A. Wyner and J. Ziv, "The rate-distortion function for source coding side information at the decoder," *IEEE Transactions on Information Theory*, vol. IT-22, no. 1, pp. 1–10, 1976.

[51] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," in *Proceedings of IEEE International Symposium on Information Theory*, Washngton, DC, June 2001, p. 294.

[52] J. N. Laneman and G. W. Wornell, "Distributed space–time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE transactions on Information Theory*, vol. 49, pp. 2415–2425, October 2003.

[53] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[54] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Transactions on Information Theory*, vol. 51, pp. 4152–4172, December 2005.

[55] P. Herhold, E. Zimmermann, and G. Fettweis, "On the performance of amplify-and-forward relay networks," in *in 5th international ITG conference on source and channel coding (SCC)*, Erlangen, Germany, January 2004.

[56] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Communication Letters*, vol. 5, no. 3, pp. 85–87, March 2001.

[57] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless network," in *Procedings of IEEE International Syposium on Information Theory*, September 2005.

[58] M. Yu, J. T. Li, and H. R. Sadjadpour, "Amplify-forward and decode-forward: The impact of location and capacity contour," in *Proceedings of the IEEE Milcom Conference*, Atlantic City, NJ, October 17-20 2005, pp. 421–425.

[59] G. M. Kraidy, N. Gresset, and J. J. Boutros, "Coding for the non-orthogonal amplify-and-forward cooperative channel," in *Proceedings of Information Theory Workshop (ITW)*, Lake Tahoe, Ca, September 2-6 2007, pp. 627–631.

[60] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the mimo amplify-and-forward cooperative channel," *submitted to IEEE transaction on Information Teory*, September 2005. [Online]. Available: http://fr.arxiv.org/pdf/cs.IT/0509006

[61] B. Schein and R. G. Gallager, "The Gaussian parallel relay network," in *International Symposium on Information Theory*, 2000, p. 22.

[62] I. Maric and R. D. Yates, "Forwarding strategies for Gaussian parallel-relay networks," in *Procedings of IEEE International Syposium on Information Theory*, 2004.

[63] A. Wittneben and B. Rankov, "Distributed antenna systems and linear relaying for gigabit MIMO wireless," in *Proceedings of IEEE Vehicular Technology Conference, VTC Fall*, Los Angeles, USA, Sept. 2004.

[64] H. Shi, T. Abe, T. Asai, and Y. Hoshino, "A relaying scheme using qr decomposition with phase control for MIMO wireless networks," in *Proc. IEEE International Conference on Communications (ICC05)*, Seoul, Korea, Sptember 2005.

[65] S. Berger and A. Wittneben, "Cooperative distributed multiuser MMSE relaying in wireless ad-hoc networks," in *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, Nov 2005.

[66] I. Hammerström, M. Kuhn, and A. Wittneben, "Distributed MIMO for cellular networks with multihop transmission protocols," in *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 2006.

[67] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, March 2006.

[68] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Low-complexity multi-user scheduling for maximizing throughput in MIMO broadcast

channel," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, Allerton, Il, October 2004.

[69] D. Gore, R. W. H. Jr., and A. J. Paulraj, "On performance of the zero-forcing receiver in presence of transmit correlation," in *IEEE International Symposium on Inforamtion Theory*, Lausanne, Switzerland, June 2003, p. 159.

[70] A. Bayesteh and A. K. Khandani, "On the user selection for MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 1086 – 1107, March 2008.

[71] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 659–672, March 2006.

[72] G. Caire and S. Shamai, "On the achievable throuput of a multiantenna gaussian channel," *IEEE Transactions on Information Theory*, vol. 49, pp. 1691–1706, July 2003.

[73] Z. Fang, Y. Hua, and J. C. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," in *Proceedings of Fourth IEEE Workshop on Sensor Array and Multichannel Processing*, July 2006.

[74] I. Hammerstroem and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Transactions on Wireless Communication*, vol. 6, no. 8, pp. 2798 – 2802, June 2007.

[75] Y. Fan, A. Adinoyi, J. Thompson, and H. Yanikomeroglu, "Antenna combining for multi-antenna multi-relay channels," *To appear in European Transactions on Telecommunications*, 2007.

[76] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Communication Magazine*, vol. 42, no. 10, pp. 74–80, October 2004.

[77] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC and beamforming for MIMO broadcast channels with many users," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 11–15, January 2007.

[78] N. J. S. Vishwanath and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of gaussian MIMO broadcast channels," *IEEE transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, October 2003.

[79] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, March 2000.

[80] ——, "Towards an information theory of large networks: an achievable rate region," *IEEE Transactions on Information Theory*, vol. 49, pp. 1877–1894, August 2003.

[81] A. F. Dana and B. Hassibi, "On the power efficiency of sensory and ad-hoc wireless networks," *IEEE Transactions on Information Theory*, 2003, submitted.

[82] S. Verdu, *Multiuser detection.* Cambridge University Press, 1998.

[83] C. E. Shannon, "Two-way communication channels," in *Proceedings of 4^{th} Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 1961, pp. 611–644.

[84] H. Sato, "The capacity of Gaussian interference channel under strong interference," *IEEE Transactions on Information Theory*, vol. 27, pp. 768–788, November 1981.

[85] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 33, pp. 710–711, September 2007.

[86] M. H. M. Costa and A. A. E. Gamal, "The capacity region of discrete memoryless interference channel with strong interference," *IEEE Transactions on Information Theory*, pp. 710–711, September 1987.

[87] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Transactions on Information Theory*, vol. 21, pp. 569–570, September 1975.

[88] ——, "Interference channels," *IEEE Transactions on Information Theory*, vol. 24, pp. 60–70, January 1975.

[89] T. M. Cover, "An achievable rate region for broadcast channel," *IEEE Transactions on Information Theory*, no. 21, pp. 399–404, 1975.

[90] H. Sato, "Two-user communication channels," *IEEE Transactions on Information Theory*, vol. 23, pp. 295–304, May 1977.

[91] A. B. Carleial, "Outer bounds on the capacity of interference channels," *IEEE Transactions on Information Theory*, vol. 29, pp. 602–606, July 1983.

[92] G. Kramer, "Outer bounds on the capacity of interference channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 581–586, March 2004.

[93] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless X networks," *submitted to IEEE Transactions on Information Theory*, 2007. [Online]. Available: http://arxiv.org/abs/0711.2824

[94] M. A. Maddah-Ali, H. Mahdavi-Doost, and A. K. Khandani, "Optimal order of decoding for max-min fairness in K-user memoryless interference channels," in *Proceedings of IEEE International Syposium on Information Theory*, Nice, France, July 2007.

[95] S. Borade, L. Zheng, , and R. Gallager, "Maximizing degrees of freedom in wireless networks," in *Proceedings of 40th Annual Allerton Conference on Communication, Control and Computing,*, October 2003, pp. 561–570,.

[96] S. Vishwanath and S. A. Jafar, "On the capacity of vector Gaussian interference channels," in *Proceedings of Information Theory Workshop (ITW)*, San Antonio, TX, October 2004, pp. 365–369.

[97] X. Shang, B. Chen, and M. J. Gans, "On the achievable sum rate for MIMO interference channels," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4313–4320, September 2006.

[98] R. S. Blum, "MIMO capacity with interference," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 793–801, June 2003.

[99] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2839–2848, November 2003.

[100] S. A. Jafar and M. Fakhreddin, "Degrees of freedom for the MIMO interference channel," in *Procedings of IEEE International Syposium on Information Theory*, Seattle, WA, July 2006.

[101] G. J. Foschini, H. Huang, K. Karakayali, R. A. Valenzuela, and S. Venkatesan, "The value of coherent base station coordination," in *Proceedings of Conference on Information Sciences and Systems (CISS)*, March 2005.

[102] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processign at the transmitting end," in *Proceeding of IEEE Vehicular Technology Conference (VTC'01)*, vol. 3, May 2001, pp. 1745–1749.

[103] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Transactions on Information Theory*, vol. 52, pp. 1522–1544, April 2006.

[104] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Decomposition of MIMO X channels," in *Proceedings of* 10^{th} *Canadian Conference on Electrical Engineering*, Edmonton, Canada, June 2007.

[105] ——, "Signaling over MIMO multi-base systems: Combination of multiple access and broadcast schemes," in *Proceedings of IEEE International Symposium on Information Theory*, Seatttle, WA, July 2006.

[106] ——, "Communication over MIMO X channels: signalling and performance analysis," Univeristy of Waterloo, Tech. Rep. UW-ECE-2007-6, 2007.

[107] S. A. Jafar, "Degrees of freedom on the MIMO X channel - the optimality of the MMK scheme," University of California, Irvine, Tech. Rep., 2006. [Online]. Available: http://arXiv:cs.IT/0607099v1

[108] S. A. Jafar and S. Shamai, "Degrees of freedom region for the MIMO X channel," 2007. [Online]. Available: arXiv.cs.IT/0607799v3

[109] V. R. Cadambe and S. A. Jafar, "Interference alignment and spatial degrees of freedom for the K user interfernce channel," *IEEE Transactions on Information Theory*, 2007, submitted.

[110] ——, "Capacity of wireless networks within $o(\log(SNR))$ - the impact of relays, feedback, cooperation and full-duplex operation," *submitted to IEEE Transactions on Information Theory*, 2008. [Online]. Available: http://arxiv.org/abs/0802.0534

[111] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: signalling and multiplexing gain," University of Waterloo, Tech. Rep. UW-ECE-2006-12, December 2006.

[112] S. A. Jafar, "Degrees of freedom of the MMK X channel - the optimality of the MMK scheme," Tech. Rep., September 2006. [Online]. Available: http://arxiv.org/abs/cs.IT/0607099

[113] L. A. Wolsey, *Integer programming.* Wiley, 1998.

APPENDIX I

Case II: $N_1 \geq M_1 > N_2 \geq M_2$

In this case, at source node $\mathcal{S}_1$ and destination node $\mathcal{R}_1$, (M_1-N_2)-dimensional sub-spaces (i.e., $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_1})$) are respectively exploited to transmit $(M_1 - N_2)$ data streams from source $\mathcal{S}_1$ to relay $\mathcal{R}_1$ and from $\mathcal{R}_1$ to $\mathcal{D}_1$ without imposing any interference at relay $\mathcal{R}_2$ and destination $\mathcal{D}_2$. For this case, we have:

$$M_1^{'} = N_2, \tag{7.1}$$

$$M_2^{'} = M_2, \tag{7.2}$$

$$N_1^{'} = N_1 + N_2 - M_1, \tag{7.3}$$

$$N_2^{'} = N_2. \tag{7.4}$$

Then $\mu^{'}_{\mathcal{S}_i\mathcal{R}_j}$ and $\mu^{'}_{\mathcal{R}_j\mathcal{D}_i}$, $\forall(i,j) \in \{1,2\}$ are chosen subjects to constraints (5.22). Consequently, we obtain the following equalities:

$$\mu_{\mathcal{S}_1\mathcal{R}_1} = \mu^{'}_{\mathcal{S}_1\mathcal{R}_1} + M_1 - N_2 \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_1} = \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} + M_1 - N_2, \tag{7.5}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_1} = \mu^{'}_{\mathcal{S}_2\mathcal{R}_1} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_1} = \mu^{'}_{\mathcal{R}_2\mathcal{D}_1}, \tag{7.6}$$

$$\mu_{\mathcal{S}_1\mathcal{R}_2} = \mu^{'}_{\mathcal{S}_1\mathcal{R}_2} \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_2} = \mu^{'}_{\mathcal{R}_1\mathcal{D}_2}, \tag{7.7}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_2} = \mu^{'}_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_2} = \mu^{'}_{\mathcal{R}_1\mathcal{D}_2}. \tag{7.8}$$

$\boldsymbol{T}_{\mathcal{S}_1}$ and $\boldsymbol{R}_{\mathcal{D}_1}$ are chosen as:

$$\boldsymbol{T}_{\mathcal{S}_1} \in \mathbb{OC}^{M_1 \times (\mu_{11}+\mu_{12})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1} = [\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1}, \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2}], \tag{7.9}$$

$$\boldsymbol{R}_{\mathcal{D}_1} \in \mathbb{OC}^{M_1 \times (\mu_{11}+\mu_{21})} \quad , \quad \boldsymbol{R}_{\mathcal{D}_1} = [\boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1}, \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1}], \tag{7.10}$$

where

$$\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \in \mathbb{OC}^{M_1 \times (N_1 - M_2)} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \in \mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2}), \tag{7.11}$$

$$\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2} \in \mathbb{OC}^{M_1 \times (\mu'_{11}+\mu_{12})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \perp \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2}, \tag{7.12}$$

$$\boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \in \mathbb{OC}^{M_1 \times (N_1 - M_2)} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \in \mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_1\mathcal{D}_2}), \tag{7.13}$$

$$\boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1} \in \mathbb{OC}^{M_1 \times (\mu'_{11}+\mu_{12})} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \perp \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1}. \tag{7.14}$$

$\boldsymbol{T}_{\mathcal{S}_2}$ and $\boldsymbol{R}_{\mathcal{R}_2}$ are chosen from $\mathbb{OC}^{M_2 \times (\mu_{21}+\mu_{22})}$ and $\mathbb{OC}^{M_2 \times (\mu_{21}+\mu_{22})}$. Furthermore, we have the following equalities:

$$\zeta_{\mathcal{S}_1\mathcal{R}_1} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_1\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{7.15}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_1} = \mu_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_2\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{7.16}$$

$$\zeta_{\mathcal{S}_1\mathcal{R}_2} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_1\mathcal{D}_2} = \mu_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}, \tag{7.17}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_2} = \mu'_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_2\mathcal{D}_2} = \mu'_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}. \tag{7.18}$$

Case III: $N_1 \geq M_1 \geq M_2 > N_2$ and $N_1 + N_2 \geq M_1 + M_2$

In this case, at $\mathcal{S}_1$ and $\mathcal{R}_1$, $(M_1 - N_2)$-dimensional sub-spaces (i.e., $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_1})$) are utilized to increase the number of data streams sent respectively from $\mathcal{S}_1$ to $\mathcal{R}_1$ and from $\mathcal{R}_1$ to $\mathcal{D}_1$ without imposing any interfer-

ence at $\mathcal{R}_2$ and $\mathcal{D}_2$ respectively. In addition at source node $\mathcal{S}_2$ and relay node $\mathcal{R}_2$, $(M_2 - N_2)$-dimensional sub-spaces, (i.e., $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_2})$) are utilized respectively to increase the number of data streams sent from $\mathcal{S}_2$ to $\mathcal{R}_1$ and from $\mathcal{R}_1$ to $\mathcal{D}_2$ by $(M_2 - N_2)$ without imposing interference at $\mathcal{R}_2$ and $\mathcal{D}_1$. We also exclude the following subspaces:

1. $(M_1 - N_2)$-dimensional sub-spaces $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_1})$ from the signaling space at $\mathcal{S}_1$ and $\mathcal{D}_1$ respectively,

2. $(M_2 - N_2)$-dimensional sub-space $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_2})$ from the signaling space at transmitter $\mathcal{S}_2$ and $\mathcal{D}_2$,

3. $(M_2 - N_1) + (M_2 - N_2)$-dimensional sub-space $\Omega\Big(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_1}\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2})\Big) \cup \Omega\Big(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_1}\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2})\Big)$ and $\Omega\Big(\boldsymbol{H}^*_{\mathcal{R}_1\mathcal{D}_1}\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_1})\Big) \cup \Omega\Big(\boldsymbol{H}^*_{\mathcal{R}_1\mathcal{D}_2}\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_2})\Big)$ from the signaling space at $\mathcal{R}_1$.

The reduced system with $(M_1^{'}, M_2^{'}, N_1^{'}, N_1^{'})$ is now defined by:

$$M_1^{'} = N_2, \tag{7.19}$$

$$M_2^{'} = N_2, \tag{7.20}$$

$$N_1^{'} = N_1 + 2N_2 - M_1 - M_2, \tag{7.21}$$

$$N_2^{'} = N_2. \tag{7.22}$$

The number of data streams are found from (5.22). Then we have:

$$\mu_{\mathcal{S}_1\mathcal{R}_1} = \mu'_{\mathcal{S}_1\mathcal{R}_1} + M_1 - N_2 \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_1} = \mu'_{\mathcal{R}_1\mathcal{D}_1} + M_1 - N_2, \tag{7.23}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_1} = \mu'_{\mathcal{S}_2\mathcal{R}_1} + M_2 - N_1 \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_1} = \mu'_{\mathcal{R}_2\mathcal{D}_1} + M_2 - N_1, \tag{7.24}$$

$$\mu_{\mathcal{S}_1\mathcal{R}_2} = \mu'_{\mathcal{S}_1\mathcal{R}_2} \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_2} = \mu'_{\mathcal{R}_1\mathcal{D}_2}, \tag{7.25}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_2} = \mu'_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_2} = \mu'_{\mathcal{R}_1\mathcal{D}_2}. \tag{7.26}$$

$\boldsymbol{T}_{\mathcal{S}_1}$ and $\boldsymbol{R}_{\mathcal{D}_1}$ are chosen as:

$$\boldsymbol{T}_{\mathcal{S}_1} \in \mathbb{OC}^{M_{\mathcal{S}_1}\times(\mu_{11}+\mu_{12})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1} = [\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1}, \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2}], \tag{7.27}$$

$$\boldsymbol{R}_{\mathcal{D}_1} \in \mathbb{OC}^{M_{\mathcal{R}_1}\times(\mu_{11}+\mu_{12})} \quad , \quad \boldsymbol{R}_{\mathcal{D}_1} = [\boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1}, \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1}], \tag{7.28}$$

where

$$\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \in \mathbb{OC}^{M_1\times(M_1-N_2)} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \in \mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2}), \tag{7.29}$$

$$\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2} \in \mathbb{OC}^{M_1\times(\mu'_{11}+\mu_{12})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \perp \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2}, \tag{7.30}$$

$$\boldsymbol{R}_{\mathcal{D}_1} \in \mathbb{OC}^{M_1\times(M_1-N_2)} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \in \mathbb{N}(\boldsymbol{H}_{\mathcal{R}_1\mathcal{D}_2}), \tag{7.31}$$

$$\boldsymbol{R}_{\mathcal{D}_1} \in \mathbb{OC}^{M_1\times(\mu'_{11}+\mu_{12})} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \perp \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1}. \tag{7.32}$$

$\boldsymbol{T}_{\mathcal{S}_2}$ and $\boldsymbol{R}_{\mathcal{D}_2}$ are chosen as:

$$\boldsymbol{T}_{\mathcal{S}_2} \in \mathbb{OC}^{M_2\times(\mu_{21}+\mu_{12})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_2} = [\boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_1}, \boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_2}], \tag{7.33}$$

$$\boldsymbol{R}_{\mathcal{D}_2} \in \mathbb{OC}^{M_2\times(\mu_{21}+\mu_{12})} \quad , \quad \boldsymbol{R}_{\mathcal{D}_2} = [\boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_2}, \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_2}], \tag{7.34}$$

where

$$\boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_1} \in \mathbb{OC}^{M_2\times(M_2-N_2)} \quad , \quad \boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_1} \in \mathbb{N}(\boldsymbol{H}_{\mathcal{S}_2\mathcal{R}_2}), \tag{7.35}$$

$$\boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_2} \in \mathbb{OC}^{M_2\times(\mu'_{21}+\mu_{22})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_1} \perp \boldsymbol{T}_{\mathcal{S}_2\mathcal{R}_2}, \tag{7.36}$$

$$\boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_2} \in \mathbb{OC}^{M_2\times(M_2-N_2)} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_2} \in \mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_2}), \tag{7.37}$$

$$\boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_2} \in \mathbb{OC}^{M_2\times(\mu'_{21}+\mu_{22})} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_2} \perp \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_2}. \tag{7.38}$$

We also have:

$$\zeta_{\mathcal{S}_1\mathcal{R}_1} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_1\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{7.39}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_1} = \mu_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_2\mathcal{D}_1} = \mu'_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{7.40}$$

$$\zeta_{\mathcal{S}_1\mathcal{R}_2} = \mu'_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_1\mathcal{D}_2} = \mu_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}, \tag{7.41}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_2} = \mu'_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \ , \ \zeta_{\mathcal{R}_2\mathcal{D}_2} = \mu'_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}. \tag{7.42}$$

Case IV: $M_1 > N_1 \geq N_2 \geq M_2$ and $N_1 + N_2 \geq M_1 + M_2$

In this case, $(M_1 - N_2)$-dimensional sub-spaces $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_2})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_2\mathcal{D}_2})$ are utilized to increase the number of data streams sent from $\mathcal{S}_1$ to $\mathcal{R}_1$ and from $\mathcal{R}_1$ to $\mathcal{D}_1$ without imposing any interference at $\mathcal{R}_2$ and $\mathcal{D}_2$. In addition $(M_1 - N_1)$-dimensional sub-space $\mathbb{N}(\boldsymbol{H}_{\mathcal{S}_1\mathcal{R}_1})$ and $\mathbb{N}(\boldsymbol{H}^*_{\mathcal{R}_1\mathcal{D}_1})$ is exploited to increase the number of data streams sent over $\mathcal{S}_1 \rightarrow \mathcal{R}_2$ and $\mathcal{R}_2 \rightarrow \mathcal{D}_1$ links,

without imposing interference at $\mathcal{R}_2$ and $\mathcal{D}_2$.

$$M_1^{'} = N_1 + N_2 - M_1, \tag{7.43}$$

$$M_2^{'} = M_2, \tag{7.44}$$

$$N_1^{'} = N_1 + N_2 - M_1, \tag{7.45}$$

$$N_2^{'} = N_1 + N_2 - M_1. \tag{7.46}$$

The number of data streams are found from (5.22). Then we have:

$$\mu_{\mathcal{S}_1\mathcal{R}_1} = \mu^{'}_{\mathcal{S}_1\mathcal{R}_1} + M_1 - N_2 \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_1} = \mu^{'}_{\mathcal{R}_1\mathcal{D}_1} + M_1 - N_2, \tag{7.47}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_1} = \mu^{'}_{\mathcal{S}_2\mathcal{R}_1} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_1} = \mu^{'}_{\mathcal{R}_2\mathcal{D}_1}, \tag{7.48}$$

$$\mu_{\mathcal{S}_1\mathcal{R}_2} = \mu^{'}_{\mathcal{S}_1\mathcal{R}_2} + M_1 - N_1 \quad , \quad \mu_{\mathcal{R}_1\mathcal{D}_2} = \mu^{'}_{\mathcal{R}_1\mathcal{D}_2} + M_1 - N_1, \tag{7.49}$$

$$\mu_{\mathcal{S}_2\mathcal{R}_2} = \mu^{'}_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \mu_{\mathcal{R}_2\mathcal{D}_2} = \mu^{'}_{\mathcal{R}_1\mathcal{D}_2}. \tag{7.50}$$

$\boldsymbol{T}_{\mathcal{S}_1}$ and $\boldsymbol{R}_{\mathcal{D}_1}$ are chosen as:

$$\boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2} \in \mathbb{OC}^{M_1 \times (\mu^{'}_{11} + \mu_{12})} \quad , \quad \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_1} \perp \boldsymbol{T}_{\mathcal{S}_1\mathcal{R}_2} \tag{7.51}$$

$$\boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \in \mathbb{OC}^{M_1 \times (M_1 - N_2 + M_1 - N_2)} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \in \mathbb{N}(\boldsymbol{H}^{*}_{\mathcal{R}_1\mathcal{D}_2}) \cup \mathbb{N}(\boldsymbol{H}^{*}_{\mathcal{R}_1\mathcal{D}_2}) \tag{7.52}$$

$$\boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1} \in \mathbb{OC}^{M_1 \times (\mu^{'}_{11} + \mu_{11})} \quad , \quad \boldsymbol{R}_{\mathcal{R}_1\mathcal{D}_1} \perp \boldsymbol{R}_{\mathcal{R}_2\mathcal{D}_1}. \tag{7.53}$$

$\boldsymbol{T}_{\mathcal{S}_2}$ and $\boldsymbol{R}_{\mathcal{D}_2}$ are chosen randomly from $\mathbb{OC}^{M_2 \times (\mu_{21} + \mu_{22})}$ and $\mathbb{OC}^{M_2 \times (\mu_{21} + \mu_{22})}$

respectively. In this case, we have:

$$\zeta_{\mathcal{S}_1\mathcal{R}_1} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \zeta_{\mathcal{R}_1\mathcal{D}_1} = \mu_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{7.54}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_1} = \mu_{\mathcal{S}_1\mathcal{R}_1} + \mu'_{\mathcal{S}_1\mathcal{R}_2} \quad , \quad \zeta_{\mathcal{R}_2\mathcal{D}_1} = \mu'_{\mathcal{R}_1\mathcal{D}_2} + \mu_{\mathcal{R}_2\mathcal{D}_2}, \tag{7.55}$$

$$\zeta_{\mathcal{S}_1\mathcal{R}_2} = \mu_{\mathcal{S}_2\mathcal{R}_1} + \mu_{\mathcal{S}_2\mathcal{R}_2} \quad , \quad \zeta_{\mathcal{R}_1\mathcal{D}_2} = \mu_{\mathcal{R}_2\mathcal{D}_1} + \mu_{\mathcal{R}_1\mathcal{D}_1}, \tag{7.56}$$

$$\zeta_{\mathcal{S}_2\mathcal{R}_2} = \mu'_{\mathcal{S}_1\mathcal{R}_1} + \mu_{\mathcal{S}_1\mathcal{R}_2} \quad , \quad \zeta_{\mathcal{R}_2\mathcal{D}_2} = \mu'_{\mathcal{R}_1\mathcal{D}_1} + \mu_{\mathcal{R}_2\mathcal{D}_1}. \tag{7.57}$$